ESTADÍSTICA

Carles Barceló i Vidal

1 ESPECIFICACIONS GENERALS 3
1.1 NOM PROGRAMA 3
1.2 AUTOR 3
1.3 TEMÀTICA 3
1.4 ASSIGNATURA(ES) 3
1.5 NIVELL ESCOLAR 3
2 INSTRUCCIONS DE FUNCIONAMENT 4
2.1 CREACIÓ I MANIPULACIÓ D'ARXIUS (Opció 1) 4
2.1.1 DISSENY DE L'ESTRUCTURA D'UN NOU ARXIU 4
2.1.2 MANIPULACIÓ D'UN ARXIU JA DISSENYAT 6
2.1.2.1 MANIPULACIÓ D'ARXIUS DADA-DADA 7
2.1.2.2 MANIPULACIÓ D'ARXIUS DIS. DE FREQÜÈNCIES . 10
2.2 TRACTAMENT ESTADÍSTIC D'UNA VARIABLE (Opció 2) . . . 11
2.2.1 ELECCIÓ DE LA VARIABLE A TRACTAR 11
2.2.2 TABULACIÓ 12
2.2.2.1 Tabulació de variables alfanumèriques . . 12
2.2.2.2 Tabulació de variables numèriques 12
2.2.3 CÀLCUL DE PARÀMETRES ESTADÍSTICS 14
2.2.4 REPRESENTACIONS GRÀFIQUES 15
2.2.5 ENREGISTRAMENT DE DADES TABULADES 15
2.3 TRACTAMENT EST. CONJUNT DE DUES VARIABLES (Opció 3). 16
2.3.1 ELECCIÓ DE LES DUES VARIABLES A TRACTAR 16
2.3.2 TABULACIÓ 16
2.3.3 CÀLCUL DE PARÀMETRES ESTADÍSTICS 18
2.3.4 RECTES DE REGRESSIÓ 19

3 ASPECTES PEDAGÒGICS 20
3.1 OBJECTIUS 20
3.2 CONEIXEMENTS PREVIS 21
3.3 FONAMENTACIÓ TEÒRICA 24
3.4 METODOLOGIA D'ÚS 27
3.4.1 Utilització puntual dels programes 27
3.4.2 Utilització en el tractament estadístic d'una enquesta 28
3.4.3 Utilització com a suport en l'ensenyament de l'estadística 28
3.4.4 Utilització com a reforç de la comprensió intuïtiva de determinats conceptes estadístics. 28
4 APÈNDIX 28
4.1 BREU DESCRIPCIÓ DE CADA ARXIU 28
4.1.1 TERCERBC.DDG 29
4.1.2 OUS1.DD2 29
4.1.3 OUS2.DD1 29
4.1.4 MEITAT.DD4 29
4.1.5 TEMPWAME.DD2 30
4.1.6 ALPARFIL.DD2 30
4.1.7 ENCRPCAS.DF1 30
4.1.8 COTXESAN.DF1 30
4.1.9 COTXAUBA.DF1 30
4.1.10 COTXAUFR.DF1 30
4.1.11 BOMBES.DF1 31
4.1.12 PARTICUL.DF1 31
4.1.13 ACCIEDAT.DF2 31
4.1.14 FISMAT.DF2 31
4.2 EXEMPLES D'EXECUCIÓ 31

1 ESPECIFICACIONS GENERALS

1.1 NOM PROGRAMA

ESTADIST.EXE

Des d'aquest programa es criden els programes ARXIUDAD.EXE, TABULAC1.EXE i TABULAC2.EXE els quals poden també cridar-se de forma autònoma.

Els acompanyen uns quants arxius estadístics creats des del programa ARXIUDAD.EXE per tal que l'usuari pugui en un principi exercitar-se en la manipulació d'aquests arxius des dels anteriors programes. En un principi s'inclouen els arxius següents:

FABRICA .DD1	TERCERBC.DDG	FISMAT .DF2	OUS1 .DD2
OUS2 .DD1	MEITAT .DD4	ENCRPCAS.DF1	ACCIEDAT.DF2
COTXESAN.DF1	COTXAUFR.DF1	COTXAUBA.DF1	PARTICUL.DF1
BOMBES .DF1	ALPARFIL.DD2	TEMPWAME.DD2	

Com es pot observar aquests arxius es caracteritzen per la seva extensió que és del tipus .DD_ o bé .DF_ .

1.2 AUTOR

Carles BARCELÓ i VIDAL (I.B. "SALVADOR ESPRIU") .

1.3 TEMÀTICA

Emmagatzematge, tabulació i anàlisi estadística elemental de petits conjunts de dades referides a una població qualsevol.

1.4 ASSIGNATURA(ES)

El programa és d'utilització preferent dins el currículum de l'assignatura de Matemàtiques. Però pot ésser emprat com a eina de càlcul, anàlisi i representació estadístics en el marc de qualsevol matèria que hagi de tabular i analitzar estadísticament petits conjunts de dades (enquestes, mostres, etc.) recollides pel professor o pels alumnes mateixos.

1.5 NIVELL ESCOLAR

B.U.P., C.O.U. i F.P.

2 INSTRUCCIONS DE FUNCIONAMENT

El conjunt està preparat per tal que els programes esmentats a l'apartat 1.1 resideixin al directori principal o en el mateix subdirectori d'un disc dur. A la boca :A se situarà el disquet que haurà de contenir els arxius de dades a tractar. En un principi s'hi poden enregistrar els arxius d'extensions .DD_ i .DF_ que acompanyen els programes. Cal tenir cura que el disquet situat a :A tingui prou espai lliure per tal d'emmagatzemar els arxius que hom vulgui crear i també per a determinades manipulacions que realitzen els diferents programes, en el decurs de les quals s'utilitza temporalment l'espai lliure del disquet.

El paquet estadístic consta de quatre programes. El principal té el nom ESTADIST.EXE i des d'aquest es carreguen i s'executen els altres tres. Les opcions que s'ofereixen són:

1. CREACIÓ I MANIPULACIÓ D'ARXIUS.
2. TRACTAMENT ESTADÍSTIC D'UNA VARIABLE.
3. TRACTAMENT ESTADÍSTIC CONJUNT DE DUES VARIABLES.

2.1 CREACIÓ I MANIPULACIÓ D'ARXIUS (Opció 1)

La primera opció és dissenyada exclusivament per a la creació i manipulació d'arxius de dades que posteriorment seran tractats estadísticament amb l'ajuda dels altres dos programes del paquet.

Abans d'introduir qualsevol conjunt de dades cal dissenyar l'estructura de l'arxiu que les contindrà. Per aquest motiu, el programa s'inicia oferint dues opcions:

1. DISSENY DE L'ESTRUCTURA D'UN NOU ARXIU.
2. MANIPULACIO D'UN ARXIU JA DISSENYAT.

2.1.1 DISSENY DE L'ESTRUCTURA D'UN NOU ARXIU

L'arxiu a dissenyar pot contenir dades no tabulades de cada un dels individus d'una població o mostra qualsevol, o bé dades ja tabulades en forma de dada-freqüència. Segons el tipus d'arxiu que hom vulgui dissenyar s'escollirà una de les dues opcions que s'ofereixen:

1. Arxiu DADA - DADA.
2. Arxiu DISTRIBUCIO de FREQUENCIES.

Ambdues opcions comencen demanant el nom amb que es coneixerà l'arxiu de dades. Aquest nom haurà de constar de 1 a 8 lletres - escrites en majúscules - o nombres, amb les mateixes restriccions que el sistema operatiu MS-DOS imposa per a la denominació dels arxius. No s'hi ha d'incloure cap tipus d'extensió ja que aquesta la determina el programa mateix en funció de la tipologia de l'arxiu.

Tot seguit el programa permet a l'usuari d'acompanyar la denominació de l'arxiu amb un breu comentari - no superior als 40 caràcters - on s'hi pot fer una descripció molt resumida del contingut de l'arxiu o de qualsevol altra característica d'aquest.

A continuació el programa demana conèixer el nombre de variables estadístiques de què constarà l'arxiu que s'està dissenyant. Si aquest és de tipus dada-dada el nombre màxim de variables permès és 12 i en el cas d'arxius dada-freqüència és 2. Els arxius dada-dada que constin de 1,2, ... o 9 variables s'enregistraran en el disquet amb el nom escollit per l'usuari seguit de l'extensió .DD1, .DD2, ... o .DD9 respectivament. Si consten de 10, 11 o 12 variables l'extensió utilitzada és .DDG . Els arxius dada-freqüència queden enregistrats amb l'extensió .DF1 o .DF2 segons constin de 1 o 2 variables.

Després de la indicació del nombre de variables de què consta l'arxiu el programa continua demanant a l'usuari que indiqui la denominació - no superior als 5 caràcters - de cada una de les variables, la seva naturalesa - numèrica (NUM) o alfanumèrica (ALF) - (veure apartat 3.3) i, si és procedent, les unitats - cm, g, ... - amb què han estat mesurades. La denominació de les unitats és limitada a 4 caràcters.

En els arxius dada-freqüència i pel que fa a les variables alfanumèriques cal indicar també els diferents "valors" que aquestes adquireixen (HOME-DONA, INSUF-SUF-BE-NOT-EXC, BLAU-MARRO-NEGRE, ...). La denominació d'aquests valors no pot superar els 5 caràcters. El nombre de valors diferents que aquestes variables poden adquirir no pot ser superior a 24, si l'arxiu consta només d'una variable (extensió .DF1), ni a 11 si consta de dues variables (extensió .DF2).

Per a les variables numèriques dels arxius dada-freqüència caldrà indicar si les dades estan o no agrupades en intervals de classe. Si ho estan, s'especificarà:

- la marca de classe del primer interval,
- la diferència constant entre marques de classe consecutives (longitud dels intervals),
- i el nombre total d'intervals de classe.

Així, per exemple, si l'arxiu ha de contenir les alçàries ja tabulades d'un grup d'alumnes que han estat organitzades en 10 intervals de classe de longitud 4 cm i on el primer interval té una marca de classe de 158 cm, caldrà introduir ordenadament els valors 158, 4 i 10.

Si les dades de la variable numèrica no estan agrupades en intervals de classe, s'especificarà:

- el valor mínim que adquireix la variable,
- la diferència mínima entre valors consecutius de la variable,
- i el nombre total de valors diferents que adquireix.

Així, per exemple, si es tracta de dissenyar un arxiu que ha de contenir la taula de freqüències dels números de sabata que calcen els alumnes d'una classe on el número mínim és un 36 i el màxim un 44, caldrà introduir ordenadament els valors 36, 1 i 9.

Una vegada dissenyada l'estructura d'un arxiu, aquesta queda enregistrada en el disc d'emmagatzematge, amb el nom escollit per l'usuari i amb l'extensió .DD_ o .DF_ que li correspongui. No convé, des del sistema operatiu, canviar el nom d'aquests arxius ja que aquests canvis només afecten el directori del disc i no el contingut dels arxius on hi figura també la denominació que l'usuari va donar a l'arxiu en el moment de dissenyar-lo.

2.1.2 <u>MANIPULACIÓ D'UN ARXIU JA DISSENYAT</u>

Aquesta opció s'inicia presentant els noms de tots els arxius d'extensió .DD_ i .DF_ que hi ha al disc d'emmagatzematge. Si al disc no hi figurés cap arxiu d'aquest tipus es produiria un error que obligaria a suspendre l'execució del programa.

Cal, tot seguit, indicar al programa quin és l'arxiu que es vol manipular. El nom caldrà introduir-lo en majúscules i especificant la seva extensió.

En funció del tipus d'arxiu que s'hagi elegit, el programa es bifurcarà en una de les dues opcions següents:

1. MANIPULACIÓ D'ARXIUS DADA-DADA.
2. MANIPULACIÓ D'ARXIUS DISTRIBUCIÓ DE FREQÜÈNCIES.

2.1.2.1 MANIPULACIÓ D'ARXIUS DADA-DADA

Aquesta opció ofereix el següent menú de possibilitats:

1. INTRODUCCIÓ DE DADES.
2. CONSULTA I/O CORRECCIÓ.
3. SUPRESSIÓ DE REGISTRES.
4. CREACIÓ DE SUB-ARXIUS.
5. LLISTATS PER IMPRESSORA.

0. FINAL DE L'OPCIÓ.

S'indica amb un asterisc (*) al davant del número aquelles opcions a les quals l'usuari té accés. Així, per exemple, si l'arxiu escollit per l'usuari no conté encara cap dada, només seran possibles les opcions 1 i 0 del menú.

L'opció 0 retorna l'execució del programa al menú inicial indicat al començament de l'apartat 2.1 . Comentarem tot seguit les altres opcions. Observeu com a l'extrem superior dret de la pantalla hi figura constantment el nom de l'arxiu que s'està manipulant.

A. INTRODUCCIÓ DE DADES (arxius .DD_)

Aquesta opció presenta a la pantalla una màscara preparada per introduir les dades de l'arxiu que s'està manipulant. A la fila superior hi figuren les denominacions de les diferents variables i a la columna esquerra els números de registre. Cada fila es correspon amb un registre que, des del punt de vista estadístic, el podem identificar amb un "individu" de la població que s'està considerant. Cada columna es correspon amb una variable o atribut. La introducció de dades caldrà fer-la indicant per a cada individu de la població el valor de cada una de les variables. D'aquesta manera, els diferents individus de la població quedaran numerats d'acord amb el número de registre que figura a la columna de l'esquerra de la pantalla.

Així, per exemple, si la població correspon als 30 alumnes d'una classe i les variables que s'estan considerant són el SEXE, el PES i l'ALÇÀRIA, caldrà anar introduint de cada alumne el valor de cada una d'aquestes tres variables.

Per acabar aquesta opció d'introducció de dades caldrà polsar la tecla preprogramada F1 seguida de la tecla ←┘.

La introducció dels diferents valors de les variables caldrà finalitzar-la amb la pulsació de la tecla ←┘. La longitud màxima d'aquests valors és de 5 caràcters, sigui quina sigui la naturalesa de la variable.

El programa no realitza cap tipus de control sobre la concordança entre la tipologia de la variable - numèrica o alfanumèrica - i el valor que l'usuari li assigna. Si la variable ha estat declarada com a numèrica i, per error, se li assigna una cadena alfanumèrica sense sentit numèric - A172, HOME, ... -, el programa l'identificarà amb el valor 0 .

Durant aquesta fase d'introducció no resulta possible, després d'haver polsat la tecla ↵, corregir un valor que acaba d'ésser introduït. Caldrà prendre nota dels errors que es puguin cometre per tal d'esmenar-los posteriorment.

És possible afegir dades a un arxiu creat amb anterioritat. En aquest cas, l'opció 1. INTRODUCCIÓ DE DADES se situa en el primer registre buit de l'arxiu.

B. CONSULTA I/O CORRECCIÓ (arxius .DD_)

Aquesta opció permet consultar i, si hom ho vol, corregir el contingut de l'arxiu que s'està manipulant. La correcció només es pot realitzar sobre la descripció que acompanya la denominació de l'arxiu, sobre el nom i unitat de cada una de les variables i sobre el valor d'aquestes en els diferents registres de l'arxiu.

A la primera "pàgina" d'aquesta opció de consulta i/o correcció, es mostra la descripció de l'arxiu i el nom, el tipus i la unitat de cada una de les variables. Si hom vol efectuar alguna correcció caldrà situar el cursor - amb l'ajuda de les tecles de moviment de cursor - davant de l'ítem a corregir, polsar la tecla ↵, escriure el nou contingut i polsar novament la tecla ↵. Acabada la correcció o consulta d'aquesta primera pàgina, es podrà prosseguir amb la consulta dels diferents registres polsant la tecla PgDn ("pàgina endavant") del teclat numèric. El programa informarà sobre el nombre de registres que conté l'arxiu i demanarà a l'usuari que indiqui a partir de quin número de registre vol iniciar la consulta i/o correcció. A continuació el programa mostrarà una pàgina amb el contingut de 15 registres de l'arxiu a partir de l'indicat per l'usuari.

La correcció dels continguts dels diferents camps dels registres de l'arxiu s'efectuarà de la mateixa forma que la indicada en el paràgraf anterior. La consulta de noves "pàgines" de l'arxiu podrà fer-se amb la pulsació de les tecles PgDn i PgUp ("pàgina enrera").

Aquesta opció de consulta i/o correcció no es donarà per acabada fins que l'usuari no polsi la tecla preprogramada F1 .

C. SUPRESSIÓ DE REGISTRES (arxius .DD_)

Si hom vol eliminar algun registre de l'arxiu de dades caldrà fer ús d'aquesta opció tot indicant al programa a partir de quin número de registre hom vol iniciar la consulta dels possibles registres a eliminar.

Com en l'opció anterior, el programa mostrarà una pàgina amb 15 registres i l'usuari indicarà, amb la pulsació de la tecla * , quins són els registres que vol eliminar. Les tecles de moviment vertical de cursor juntament amb les PgDn i PgUp facilitaran a l'usuari la localització dels registres a eliminar. La pulsació de la tecla F1 donarà per acabada aquesta d'opció.

Sempre que es faci ús d'aquesta opció cal comprovar amb anterioritat que el disc d'emmagatzematge conté lliure prou espai per tal d'enregistrar temporalment un arxiu - de grandària similar al que s'està manipulant - i que el programa crea en el decurs d'aquesta opció de supressió de registres. Si no hi hagués prou espai en el disc es produiria un error que interrompria l'execució del programa.

D. CREACIÓ DE SUB-ARXIUS (arxius .DD_)

Aquesta opció només s'ofereix en el cas que s'estigui manipulant un arxiu dada-dada que contingui dos o més variables. Està pensada per a facilitar la creació - a partir d'arxius contenint moltes variables d'una població - d'arxius més reduïts albergant només una o més de les variables de l'arxiu "mare". El nombre de registres del sub-arxiu és el mateix que el de l'arxiu original; varia només el nombre de variables.

Per a l'execució d'aquesta opció l'usuari indicarà primerament el nombre de variables que ha de contenir el sub-arxiu i posteriorment assenyalarà amb la pulsació de la tecla * quines són les variables que ha de contenir. Les tecles de moviment vertical de cursor permetran indicar amb facilitat quines són aquestes variables. A continuació el programa reclamarà de l'usuari la introducció del nom del sub-arxiu i el comentari que l'ha d'acompanyar. Són d'aplicació els mateixos comentaris que els indicats al començament de l'apartat 2.1.1 dedicat al disseny de nous arxius.

E. LLISTATS PER IMPRESSORA (arxius .DD_)

Aquesta opció permet obtenir per impressora una descripció complerta de les característiques de l'arxiu que s'està manipulant i dels continguts dels registres de dades (vegeu mostres de llistats als apèndixs finals).

2.1.2.2 <u>MANIPULACIÓ D'ARXIUS DISTRIBUCIÓ DE FREQÜÈNCIES</u>

Aquesta opció ofereix el següent menú de possibilitats:

1. INTRODUCCIÓ DE DADES.
2. CONSULTA I/O CORRECCIÓ.
3. LLISTATS PER IMPRESSORA.

0. FINAL DE L'OPCIÓ.

Si l'arxiu no conté encara cap dada - tan sols ha sigut dissenyat -, només seran possibles les opcions 1 i 0 del menú, les quals vindran indicades amb un asterisc (*).

La introducció de dades cal realitzar-la d'una sola vegada. Per aquest motiu, si aquelles haguéssin estat introduïdes amb anterioritat, el menú només assenyalaria com possibles les opcions 2, 3 i 0.

A. INTRODUCCIÓ DE DADES (arxius .DF_)

En els arxius de distribució de freqüències, les dades a introduir no són pròpiament els valors de la variable o variables de la població - que ja varen ésser introduïts en la fase de disseny - sinó les freqüències d'aquests valors. Cal, doncs, introduir les freqüències absolutes que figuren en una taula de freqüències (cas .DF1) o en una taula de contingència (cas .DF2).

En el cas d'arxius del tipus .DF1 el programa mostra ordenadament els diferents valors o marques de classe de la variable i l'usuari ha d'anar introduint les corresponents freqüències. La introducció de la freqüència de l'últim ítem finalitza automàticament aquesta opció d'introducció de dades.

El el cas d'arxius del tipus .DF2 el programa presenta una taula de contingència "buida" on hi figuren els valors o marques de classe de les dues variables que s'estan considerant. L'usuari ha d'anar introduint ordenadament en cada casella les freqüències corresponents. Un cop introduïda l'última de les freqüències el programa no dóna per finalitzada l'opció d'introducció de dades fins que l'usuari no polsi l'espaiador. D'aquesta manera es facilita la possibilitat d'obtenir una còpia escrita de la taula de contingència tot fent un "abocat de pantalla" amb la pulsació de la tecla PrtSc.

C. LLISTATS PER IMPRESSORA (arxius .DF_)

En els llistats per impressora que aquesta opció proporciona s'expliciten les característiques de l'arxiu que s'està manipulant així com els diferents valors o marques de classe de la variable o variables de què es compon, acompanyats de les seves respectives freqüències. En el cas d'arxius .DF2 els llistats no es presenten en forma de taula de doble entrada sinó en forma lineal.

2.2 TRACTAMENT ESTADÍSTIC D'UNA VARIABLE (Opció 2)

Aquesta segona opció del programa principal està pensada per a facilitar la tabulació de dades, càlcul de paràmetres estadístics i obtenció de representacions gràfiques, amb referència a una variable estadística qualsevol que ha estat suposadament mesurada sobre els "individus" d'una determinada "població". Permet d'utilitzar igualment dades ja tabulades. Totes les dades a utilitzar en aquesta opció hauran d'haver estat prèviament enregistrades en el disc d'emmagatzemament a través de l'opció 1 de CREACIÓ I MANIPULACIÓ DE DADES .

El programa ofereix les opcions següents:

1. ELECCIÓ DE LA VARIABLE A TRACTAR.
2. TABULACIÓ.
3. CÀLCUL DE PARÀMETRES ESTADÍSTICS.
4. REPRESENTACIONS GRÀFIQUES.
5. ENREGISTRAMENT DE DADES TABULADES.

0. FINAL DE L'OPCIÓ.

En el moment d'iniciar aquesta opció, només són vàlides les opcions 1 i 0 que apareixen precedides del signe * .

Observeu com en l'anterior menú no hi figura cap opció de LLISTATS PER IMPRESSORA ja que s'ha suposat que l'usuari pot fer ús de la tecla PrtSc per tal d'obtenir els "abocats de pantalla" que l'interessin més.

2.2.1 ELECCIÓ DE LA VARIABLE A TRACTAR

Aquesta opció presenta una relació de tots els arxius que tenen una extensió del tipus .DD_ o .DF_ i que figuren en el disc d'emmagatzematge situat en la boca :A . L'usuari haurà d'indicar el nom de l'arxiu que conté la variable que vol

tractar. El nom de l'arxiu caldrà escriure'l acompanyat de la corresponent extensió que haurà de ser del tipus .DD_ o bé la .DF1 . Els arxius .DF2 (taules de contingència) no poden ésser tractats des d'aquesta segona opció del programa principal.

A continuació el programa presenta una relació de les característiques de l'arxiu que ha estat escollit: nom, breu descripció, nombre de registres que conté i relació de variables (amb indicació del seu nom, tipus i unitat de mesura). Si l'arxiu és del tipus .DD_ i conté més d'una variable, l'usuari haurà d'assenyalar amb un * quina és la variable que vol tractar estadísticament. La pulsació de les tecles de moviment vertical del cursor permetran situar el cursor sobre la variable escollida.

L'opció finalitza amb la pulsació de l'espaiador fent que l'execució del programa retorni a la presentació del menú inicial que inclou ja, com a nova opció, la que fa referència a la TABULACIÓ de la variable escollida.

2.2.2 TABULACIÓ

2.2.2.1 Tabulació de variables alfanumèriques

Si la variable a tabular és alfanumèrica, el programa comença per fer un inventari dels diferents valors que pren la variable en el conjunt de dades enregistrades en l'arxiu. Tot seguit informa l'usuari sobre el nombre de dades que conté l'arxiu i sobre el nombre de valors diferents que pren la variable estadística, tot especificant quins són aquests valors i quines les seves respectives freqüències d'aparició.

En previsió que la variable alfanumèrica que s'estigui tabulant sigui de tipus ORDINAL , el programa ofereix a l'usuari la possibilitat d'ordenar com ell vulgui els diferents valors que la variable ha pres. Cada un d'aquests valors s'identifica amb una lletra A, B, C, ... de l'abecedari i cal només introduir ordenadament les lletres corresponents a cada valor, tot seguint l'ordre que l'usuari hagi determinat prèviament.

A continuació el programa presenta una taula de freqüències tot indicant per a cada valor de la variable la seva freqüència absoluta i percentual. La pulsació de l'espaiador retorna l'execució del programa a la presentació del menú inicial.

2.2.2.2 Tabulació de variables numèriques

Si la variable que s'està tabulant és de tipus numèric i procedent d'un arxiu tipus .DD_ , el programa comença buscant en l'arxiu quins són els valors màxim i mínim que pren la variable i

quina és la diferència mínima que s'observa entre dades consecutives no coincidents. A continuació el programa demana confirmació a l'usuari sobre si aquesta diferència mínima entre dades que ha estat detectada coincideix efectivament amb la diferència mínima real entre valors qualssevol de la variable que s'està tabulant.

Així, per exemple, podria ocórrer que el programa, tot tabulant les alçàries en cm dels alumnes d'una classe, detectés una diferència mínima de 2 cm entre el conjunt de dades enregistrades en l'arxiu quan, realment, les mesures havien estat realitzades amb una precisió d'1 cm. En aquest cas, l'usuari haurà d'informar el programa que la diferència mínima real entre dades és d'1 cm.

Aquesta puntualització sobre la diferència mínima entre dades, el programa la utilitza a l'hora de decidir amb quina precisió s'han de presentar els valors dels diferents paràmetres estadístics, i també en el tractament que realitza de les variables numèriques contínues no agrupades en intervals de classe. Per a més informació sobre aquest punt cal remetre's a l'apartat 3.3 d'aquesta documentació.

A continuació el programa interroga l'usuari sobre quin tractament - DISCRET o CONTINU - vol donar a la variable que s'està tabulant. No és tampoc ara el moment de puntualitzar sobre les diferències entre ambdós tractaments. Remetem el lector a la lectura de l'apartat 3.3 d'aquesta documentació.

Si s'ha decidit considerar la variable numèrica com a discreta, el programa demana a l'usuari si vol o no agrupar les dades en intervals de classe d'igual longitud. En cas afirmatiu, es pregunta la longitud comuna que hom vol que tinguin els intervals. Aquesta longitud haurà d'ésser múltiple de la diferència mínima entre dades. El primer interval de classe el determina el programa mateix en funció de la longitud dels intervals que s'hagi triat i de la diferència mínima entre dades.

Si el tractament que hom vol donar a la variable és continu, el programa ja dóna per descomptat que les dades s'hauran d'agrupar en intervals. Per aquest motiu passa directament a preguntar a l'usuari la longitud comuna que han de tenir els intervals de classe i el valor de l'extrem inferior del primer d'aquests intervals. La longitud introduïda no podrà ésser inferior a la diferència mínima entre dades.

El programa no contempla la possibilitat d'agrupar les dades en intervals que no siguin de la mateixa longitud.

Tant si la variable ha estat considerada discreta com contínua, en el cas que la longitud dels intervals coincideixi amb la diferència mínima entre dades, el programa informa que aquesta coincidència fa que es considerin les dades com NO AGRUPADES.

Si la variable numèrica que s'està tractant procedeix d'un arxiu tipus .DF1 el programa no interroga l'usuari sobre cap aspecte relacionat amb la diferència mínima entre dades, i es limita només a preguntar sobre el tipus de tractament que hom vol donar a la variable. Si la variable té les dades ja agrupades en intervals de classe, el programa es limita a informar sobre el nombre d'intervals i sobre la longitud d'aquests. En cas que les dades no estiguéssin agrupades, el programa procediria tal com s'ha explicat en els paràgrafs anteriors.

L'opció finalitza presentant una taula on s'expliciten per a cada un dels intervals les corresponents marques de classe, les freqüències absolutes, les freqüències percentuals, les freqüències absolutes acumulades i les freqüències acumulades percentuals. La pulsació de l'espaiador retorna al menú principal de l'opció.

2.2.3 CÀLCUL DE PARÀMETRES ESTADÍSTICS

En el cas que la variable sigui alfanumèrica el programa mostra només el valor de la moda o modes de la distribució estadística.

Si la variable és numèrica el programa presenta:

- Paràmetres de centralització: moda/es, mediana i mitjana.

- Paràmetres de dispersió: valor màxim, valor mínim, rang, primer quartil, tercer quartil, rang interquartílic, desviació típica, tant per cent de dades compreses entre $M-\sigma$ i $M+\sigma$ (essent M la mitjana i σ la desviació típica) i tant per cent de dades compreses entre M-2 i M+2 .

A més a més, el programa, després de presentar els paràmetres de centralització, il.lustra gràficament la posició de la mediana i la mitjana amb referència als valors extrems de la distribució.

Igualment, el programa proporciona una interpretació gràfica de la posició del tres quartils amb referència als valors màxims i mínims de la distribució.

L'opció finalitza il.lustrant, de forma similar a les anteriors, la posició dels punts $M-\sigma$ i $M+\sigma$.

2.2.4 <u>REPRESENTACIONS GRÀFIQUES</u>

Si la variable és alfanumèrica el programa es limita a representar gràficament la distribució de freqüències en forma de diagrama de sectors oferint a l'usuari la possibilitat d'escollir entre un cercle o un semicercle com a suport del diagrama.

En el cas de variables numèriques contínues el programa representa el corresponent histograma de freqüències i a continuació l'histograma de les freqüències acumulades. Ambdós diagrames s'acompanyen de dues escales, una indicant les freqüències absolutes i l'altra les freqüències percentuals. Després de cada diagrama es dibuixa el corresponent polígon de freqüències.

Si la variable que s'està tractant és numèrica de tipus discret, l'opció s'inicia demanant a l'usuari si vol utilitzar diagrames de barres o bé histogrames en les representacions dels diagrames de freqüències. La dinàmica subsegüent coincideix amb l'anterior amb l'única diferència que no es representen els polígons de freqüències en el cas que l'usuari hagi optat pels diagrames de barres.

2.2.5 <u>ENREGISTRAMENT DE DADES TABULADES</u>

Aquesta opció només s'ofereix en el cas que s'estigui manipulant una variable procedent d'un arxiu dada-dada.

Està dissenyada per a possibilitar l'enregistrament en forma d'arxiu del tipus .DF1 de la taula de freqüències de la variable que s'ha estat tractant i que procedeix d'un arxiu del tipus .DD_ . L'arxiu que d'aquesta manera es crea - el qual queda enregistrat en el disquet d'emmagatzematge - podrà ésser posteriorment tractat des de la mateixa opció del programa principal. Conté el mateix tipus d'informació que l'indicat en el apartat 2.1.1 (arxius dada-freqüència) d'aquesta documentació, acompanyada de les corresponents freqüències, calculades en l'opció de TABULACIÓ.

Aquesta opció es limita a demanar a l'usuari el nom amb què vol designar l'arxiu de dades tabulades i la descripció que l'ha d'acompanyar. El programa controla l'existència en el disc d'emmagatzemamemt d'algun possible arxiu amb el mateix nom i amb la mateixa extensió (.DF1). En cas que hi fos, el programa

demanaria a l'usuari la confirmació del nom de l'arxiu prèviament escollit. L'enregistrament d'un arxiu amb el mateix nom que un altre resident en el disc, provoca la pèrdua d'aquest últim.

2.3 TRACTAMENT ESTADÍSTIC CONJUNT DE DUES VARIABLES (Opció 3)

Aquesta tercera opció del programa principal persegueix com objectiu principal la tabulació conjunta de dues variables mesurades sobre els mateixos individus d'una població, i el càlcul de diferents paràmetres estadístics que ajudin a estudiar la possible correlació entre ambdues variables.

L'opció ofereix d'entrada el menú següent:

1. ELECCIÓ DE LES DUES VARIABLES A TRACTAR.
2. TABULACIÓ.
3. CÀLCUL DE PARÀMETRES ESTADÍSTICS.
4. RECTES DE REGRESSIÓ
5. ENREGISTRAMENT DE DADES TABULADES.

0. FINAL DE L'OPCIÓ.

D'entrada només són possibles les opcions 1 i 0 que apareixen precedides del signe * .

2.3.1 ELECCIÓ DE LES DUES VARIABLES A TRACTAR

Aquesta opció es desenvolupa de forma similar a la indicada en l'apartat 2.2.1 . Només són vàlids els arxius d'extensió .DF2 o bé els arxius .DD2, .DD3, ...DDG .

Si s'escull un arxiu del tipus dada-dada contenint més de dues variables, caldrà, com de costum, assenyalar amb un * les dues variables que hom desitja estudiar.

2.3.2 TABULACIÓ

En el cas de variables procedents d'arxius .DD_ , la fase inicial d'aquesta opció de tabulació conjunta d'ambdues variables ve a ser la repetició dues vegades del procés explicat en l'apartat 2.2.2. Esmentar només que les limitacions del programa fan que, si hom decideix agrupar en intervals de classe els valors que adquireix una variable numèrica, el nombre d'aquests no pot ésser superior a 11. Igualment, les variables numèriques no agrupades en intervals i les variables alfanumèriques veuen limitat a 11 el nombre màxim de valors diferents que poden adquirir.

Si les dues variables seleccionades procedeixen d'un arxiu del tipus .DF2, aquesta opció es limita, en un principi, a informar sobre les característiques de cada una de les variables: nom, unitat de mesura, tipus i nombre de valors que adquireix o nombre d'intervals en què està agrupada.

A continuació el programa ofereix a l'usuari la possibilitat de visualitzar quatre tipus de taules de contingència:

1. TAULA DE FREQÜÈNCIES ABSOLUTES.
2. TAULA DE FREQÜÈNCIES PERCENTUALS.
3. TAULA DE FREQÜÈNCIES PERCENTUALS PER FILES.
4. TAULA DE FREQÜÈNCIES PERCENTUALS PER COLUMNES.

La primera taula mostra la freqüència absoluta de cada una de les caselles de la taula de contingència, situades en les interseccions de les columnes corresponents als diferents valors o marques de classe de la primera variable i de les files lligades als valors o marques de classe de la segona variable. Presenta igualment les freqüències marginals de cada una de les dues variables de la taula.

El programa pren com a primera variable (variable X o variable columnes) la que l'usuari va escollir en primer lloc en el moment de l'elecció de variables, o, si no ha tingut oportunitat d'escollir-les, la que ocupa el primer lloc en la relació de variables de l'arxiu.

La segona taula presenta, per a cada una de les caselles, la freqüència percentual referida al total de dades - individus - de la població. S'acompanya igualment de les freqüències percentuals marginals.

La tercera taula presenta la freqüència percentual de cada casella referida a la freqüència absoluta marginal de la columna a la qual pertany. Si la freqüència marginal d'una columna és zero, no es calculen les freqüències percentuals d'aquella columna.

La quarta taula mostra la freqüència percentual de cada casella amb referència a les corresponents freqüències marginals de la segona variable (files). Si alguna d'aquestes és zero, deixen de calcular-se les freqüències percentuals de la fila corresponent.

2.3.3 <u>CÀLCUL DE PARÀMETRES ESTADÍSTICS</u>

Aquesta opció ofereix d'entrada el següent menú:

1. M i σ DE LA PRIMERA VARIABLE (COLUMNES).
2. M i σ DE LA SEGONA VARIABLE (FILES).
3. COEFICIENT DE CORRELACIÓ (PEARSON).
4. COEFICIENT DE CORRELACIÓ PER RANGS (SPEARMAN).
5. M i σ DE CADA COLUMNA.
6. M i σ DE CADA FILA.

0. FINAL DE L'OPCIÓ.

Les opcions 1 i 6 tindran sentit només en el cas que la primera variable sigui de tipus numèric. De la mateixa manera, si la segona variable és numèrica, resultaran vàlides les opcions 2 i 5 del menú. L'opció 3 només tindrà sentit en el cas que ambdues variables siguin numèriques. En qualsevol cas, però, el programa mostra un * al davant dels números de les opcions vàlides.

Les opcions 1, 2 i 3 es limiten senzillament a mostrar els valors dels paràmetres corresponents.

L'opció 4 de càlcul del coeficient de Spearman s'ofereix, d'entrada, sigui quina sigui la parella de variables que s'estigui tractant sempre que no n'hi hagi cap que sigui de tipus alfanumèric i amb un nombre de dades superior a 1000. En el cas que una de les variables sigui alfanumèrica l'opció adverteix a l'usuari que el coeficient de Spearman que es calcularà només tindrà sentit si la variable o variables alfanumèriques que s'estan tractant són de tipus ORDINAL, és a dir, si existeix un ordre implícit en els valors que prenen (p.e., INSUF < SUF < BE < NOT < EXC). Hom ha de suposar, a més a més, que aquestes variables han estat tabulades tenint en compte aquest ordre implícit. Suposant que l'usuari confirmi al programa que efectivament concorren aquestes condiciones per al càlcul del coeficient de Spearman, el programa inicia el procés de càlcul. Si el nombre de dades és elevat, aquest procés pot perllongar-se durant uns pocs minuts.

L'opció 5 ens mostra la mitjana M i la desviació típica σ de les distribucions marginals de la segona variable (files) per a cada un dels valors o intervals de la primera variable (columnes). Els anteriors paràmetres no es calculen en aquelles columnes en les quals la freqüència marginal corresponent val 0 .

L'opció 6 es similar a l'anterior, canviant només el paper que juguen ambdues variables.

2.3.4 RECTES DE REGRESSIÓ

Aquesta opció només s'ofereix en el cas que ambdues variables siguin numèriques. S'inicia informant l'usuari sobre les característiques d'ambdues variables - denominació, unitat de mesura, valor màxim, valor mínim, ... - i demanant-li que trii quines unitats voldrà prendre sobre cada un dels dos eixos de coordenades respecte dels quals es representaran les rectes de regressió. D'entrada s'aconsella prendre com a unitat damunt de cada eix la dècima part del rang de la variable corresponent.

A continuació el programa dibuixa els dos eixos de coordenades, graduats en correspondència amb les marques de classe o amb els valors de cada variable que han estat considerats en el procés de tabulació, i tot seguit representa el núvol de punts de la distribució conjunta. La pulsació de la tecla F6 ha de permetre a l'usuari canviar les unitats de representació si no queda satisfet amb la distribució de punts obtinguda. Si la graduació de l'eix d'ordenades queda molt petita amb referència a la de l'eix d'abscisses (el núvol de punts queda molt concentrat al costat d'aquest eix) caldrà disminuir el valor assignat a la unitat sobre l'eix d'ordenades respecte a la considerada sobre l'eix d'abscisses. Es procedirà a l'inrevés en cas que la graduació sobre l'eix d'abscisses resulti molt petita amb referència a la de l'eix d'ordenades.

Una vegada aconseguida una graduació correcta dels eixos, la pulsació de l'espaiador farà que el programa presenti el següent menú d'opcions:

1. RECTA DE REGRESSIÓ DE Y SOBRE X .
2. RECTA DE REGRESSIÓ DE X SOBRE Y .
3. DISTRIBUCIÓ DE Y DINS DE CADA INTERVAL DE X .
4. DISTRIBUCIÓ DE X DINS DE CADA INTERVAL DE Y .
5. CANVI D'UNITATS SOBRE ELS EIXOS.

0. FINAL DE L'OPCIÓ.

L'opció 1 representa el núvol de punts juntament amb la recta de regressió de Y sobre X . Per tal de referenciar la posició dels punts de la distribució estadística respecte a la mitjana de cada una de les variables, hom representa també un sistema de coordenades que té per origen el centre de gravetat de la distribució, és a dir, el punt d'abscissa igual a la mitjana de la variable X i d'ordenada igual a la mitjana de la variable Y . La pulsació de la tecla F6 permet visualitzar també la recta de regressió de X sobre Y .

L'opció 2 actua de forma similar a l'opció 1 representant la recta de regressió de X sobre Y .

L'opció 3 representa, per a cada un dels intervals de la variable X, les corresponents mitjanes de la distribució de la variable Y dins d'aquests intervals, dibuixant a més a més sengles entorns centrats en aquestes mitjanes i de radis iguals a les respectives desviacions típiques de la variable Y dins de cada un d'aquests intervals. Una poligonal uneix els successius centres d'aquests entorns. Aquesta opció només té sentit si es treballa amb un volum de dades certament nombrós ja que sinó, la presència d'intervals de molt poca freqüència - 0, 1, 2, ... - desvirtua totalment la representació gràfica.

L'opció 4 realitza el mateix que l'opció 3 intercanviant només el paper de les dues variables.

Finalment, l'opció 5 permet canviar novament les unitats de representació sobre els eixos de coordenades.

Cal comentar que el núvol de punts no es representa quan el nombre d'aquests és superior a 1000 donat que la seva representació resultaria certament molt confusa. En aquests casos les opcions 1 i 2 es limiten a representar les respectives rectes de regressió.

Quan les variables estadístiques amb les quals s'opera procedeixen d'arxius tipus .DF2 , el fet de desconèixer quina és la correspondència real entre els valors de la variable X i els de la variable Y , fa que el núvol de punts es representi escollint a l'atzar els valors d'ambdues variables dins de cada un dels respectius intervals de classe. Així, per exemple, si se sap que hi ha 10 individus de la població que tenen una estatura compresa entre 169.5 i 172.5 cm i un pes entre 64.5 i 68.5 kg, el programa escollirà aleatòriament 10 valors dins l'interval [169.5,172.5) que farà correspondre amb 10 valors dins l'interval [64.5,68.5), també triats aleatòriament. Tot seguit el programa representarà els punts corresponents. L'elecció aleatòria dels valors la realitza el programa de manera que s'obtingui sempre el mateix núvol de punts cada vegada que s'estudiï la mateixa distribució conjunta de variables.

3 ASPECTES PEDAGÒGICS

3.1 OBJECTIUS

Entre d'altres, els objectius que aquest conjunt de programes persegueix són, fonamentalment, els següents:

- Facilitar l'àgil manipulació, càlcul, representació gràfica i interpretació estadística de conjunts de dades més nombrosos que els que habitualment es poden utilitzar en el marc escolar actual.

- Ajudar a adquirir i augmentar la intuició al voltant de certs conceptes i paràmetres estadístics - correlació i regressió lineals, mediana i mitjana, desviació típica, ... - que molt sovint són considerats únicament des d'un punt de vista numèric formulístic en no poder disposar de mitjans que permetin una àgil i ràpida interpretació gràfica i numèrica.

Aquest paquet de programes no ha estat concebut en cap moment com una potent eina de càlcul estadístic capaç de suportar qualsevol tipus de dades. El seu disseny ve marcat per una clara intenció pedagògica d'incorporar l'ensenyament i el pensament estadístics en els nostres currículums escolars. Cal tenir cura, doncs, del tipus i de la magnitud de les dades a manipular amb aquests programes per tal de no depassar les seves possibilitats reals.

3.2 <u>CONEIXEMENTS PREVIS</u>

És convenient que l'usuari d'aquests programes tingui un coneixement mínim dels termes i conceptes que s'hi utilitzen per què, com ja ho hem indicat, no és un objectiu dels programes ensenyar a l'usuari la forma com es calculen els diferents paràmetres estadístics sinó ajudar-lo a interpretar-los correctament. Entenem, però, que l'aprenentatge dels conceptes estadístics involucrats en aquest paquet de programes pot realitzar-se simultàniament amb el desenvolupament progressiu d'aquest "software" per tal d'anar reforçant el procés d'aprenentatge.

Més concretament, els programes d'aquest paquet estan relacionats de forma directa o indirecta amb els següents conceptes i tècniques de l'ensenyament de l'estadística:

A RECOLLIDA DE DADES

A.1 Utilització d'un cens

(Per cens cal entendre qualsevol informació que s'obté de tots els elements d'una població)

A.1.1 Recompte simple de dades en poblacions petites, sense problemes de significació, mesura o classificació d'aquestes (p.e., "obtenció de dades relatives a tots els alumnes d'una classe").

A.1.2 Problemes relatius a la interpretació i obtenció de dades procedents de poblacions petites i a la utilització de qüestionaris per a la realització d'enquestes d'opinió.

A.1.3 Sobre poblacions extenses. Problemes derivats del procés de distribució i recollida de dades, de la no resposta a determinades preguntes dels qüestionaris, de l'ambigüitat en la formulació de certes preguntes, del control d'errors, etc.

A.2 Tipus de dades.

A.2.1 Dades discretes: escales nominal, ordinal i de raó.
A.2.2 Dades contínues: escales d'interval (p.e., la temperatura) i de raó (p.e., el pes).
A.2.3 Agrupació de dades: intervals de classe, pèrdua de precisió.
A.2.4 Dades aparellades: variables no numèriques i variables numèriques discretes.
A.2.5 Dades aparellades: variables numèriques contínues (agrupades i no agrupades).

A.3 Recollida de dades.

A.3.1 Recollida de dades pròpies: precisió en els recomptes i mesures, ambigüitat de les definicions, problemes de versemblança d'algunes respostes, problemes derivats de la "no resposta" de certes preguntes.
A.3.2 Utilització de dades ja tabulades: problemes relacionats amb les fonts subministradores de les dades i amb la precisió i la significació d'aquestes.
A.3.3 Utilització de qüestionaris d'elaboració pròpia per a la recollida de dades: problemes relacionats amb la redacció de preguntes, amb la formulació de preguntes "tendencioses", amb les definicions de determinats conceptes, amb el control de la versemblança de les respostes, amb el tractament a donar a les preguntes "no contestades", amb la utilització o no de les tècniques d'entrevistes personals.

B. TABULACIÓ I REPRESENTACIÓ DE DADES

B.1 Tabulació

B.1.1 Construcció, per recompte directe, de taules de freqüència.
B.1.2 Construcció, per recompte directe, de taules de freqüències de doble entrada (taules de contingència).
B.1.3 Reducció de taules: dades significatives, contraposició entre facilitat de lectura i pèrdua d'informació.

B.2 Representacions gràfiques.

- B.2.1 Diagrames de barres: variables no numèriques (escales ordinal i nominal) i numèriques (discretes).
- B.2.2 Pictogrames.
- B.2.3 Diagrames de sectors: radi constant.
- B.2.4 Diagrames de sectors: àrea proporcional a l'àrea total.
- B.2.5 Diagrames de barres: variables numèriques contínues (dades agrupades en intervals d'igual longitud).
- B.2.6 Histogrames: variables numèriques discretes i contínues agrupades en intervals de classe d'igual longitud. Problemes de frontera.
- B.2.7 Histogrames: variables numèriques discretes i contínues agrupades en intervals de classe de desigual longitud.
- B.2.8 Diagrames de freqüències acumulades: variables discretes i contínues agrupades.
- B.2.9 Diagrames de dispersió.

C PARÀMETRES ESTADÍSTICS

C.1 Paràmetres de posició.

- C.1.1 Moda d'un conjunt de dades no agrupades.
- C.1.2 Mediana d'un conjunt de dades no agrupades.
- C.1.3 Mitjana d'un conjunt de dades no agrupades.
- C.1.4 Mitjana: canvi d'origen en el cas de dades no agrupades.
- C.1.5 Classe modal. Efectes provocats per canvis en la llargada dels intervals de classe.
- C.1.6 Mitjana d'un conjunt de dades agrupades, sense cap tipus de transformació.
- C.1.7 Mitjana d'un conjunt de dades agrupades: canvi lineal d'escala.
- C.1.8 Mediana d'un conjunt de dades agrupades utilitzant freqüències acumulades. Distinció dels casos discret i continu.

C.2 Paràmetres de dispersió.

- C.2.1 Rang d'un conjunt de dades no agrupades.
- C.2.2 Quartils d'un conjunt de dades no agrupades.
- C.2.3 Quartils d'un conjunt de dades agrupades de tipus discret mitjançant la utilització de les freqüències acumulades.
- C.2.4 Quartils d'un conjunt de dades agrupades de tipus continu fent ús de les freqüències acumulades.
- C.2.5 Rang interquartílic.
- C.2.6 Variança i dispersió típica d'un conjunt de dades no agrupades.

C.2.7 Variança i dispersió típica d'un conjunt de dades agrupades en intervals.
C.2.8 Variança i dispersió típica: canvis lineals d'escala.
C.2.9 Relació entre percentils i dispersió típica.

C.3 Regressió i correlació.

C.3.1 Ajust d'una distribució conjunta de dues variables numèriques al voltant d'una recta adequada.
C.3.2 Corbes d'ajust: dibuix a cop d'ull.
C.3.4 Càlcul de la recta de regressió pel mètode de mínims quadrats.
C.3.5 Diagrames de dispersió i dependència entre dues variables.
C.3.6 Mesura de la dependència entre dues variables: cas no paràmetric. Coeficient de correlació de Spearman.
C.3.7 Corbes d'ajust: la seva utilització en estimacions.
C.3.8 Coeficient de correlació: càlcul i interpretació.

3.3 FONAMENTACIÓ TEÒRICA

No és ara el moment d'explicar els conceptes estadístics involucrats en el programa ja que es poden consultar en qualsevol manual d'estadística elemental.

Entenem, però, que resulta necessari puntualitzar la distinció que el programa fa entre els diferents tipus de variables ja que aquella podria ésser font d'algunes confusions.

D'entrada, el primer programa dedicat a la creació i manipulació d'arxius demana a l'usuari la distinció entre dos tipus de variables: numèriques i alfanumèriques. Si bé potser la denominació de variables alfanumèriques pot semblar poc afortunada, s'ha cregut oportú adoptar-lo a semblança de la classificació de variables que realitza el llenguatge BASIC. Com a tal, s'ha d'entendre qualsevol variable que prengui valors no subjectes d'aplicar-los manipulacions de tipus algebraic.

Els altres dos programes dedicats al tractament estadístic d'una variable o bé al tractament conjunt de dues variables, a l'hora de manipular variables de tipus numèric demanen a l'usuari informació sobre el tipus de tractament - DISCRET o CONTINU - que cal donar a la variable que s'està considerant.

El tractament discret d'una variable numèrica, el programa el realitza identificant totes les dades pertanyents a un mateix interval de classe amb el valor de la corresponent marca de classe. Mentre que si la variable és considerada de tipus continu, el programa considera les dades contingudes dins un

mateix interval de classe uniformement repartides dins d'aquest. Així, per exemple, si la freqüència absoluta del valor d'una variable sobre un interval [39.5,42.5) val 6, un tractament discret identificarà els 6 valors de la variable amb la marca de classe 41, mentre que un tractament continu suposarà que els 4 valors són 39.75 , 40.25 , 40.75 , 41.25 , 41.75 i 42.25 :

```
        39.75   40.25   40.75   41.25   41.75   42.25

  39.5 .   .       .       .       .       .       .   .  42.5
```

Aquesta distinció no afecta a la mitjana de la distribució però sí a la resta de paràmetres:

a. Valors màxim i mínim. Rang.

Si el valor mínim d'un conjunt de dades és, p.e., 154 i una vegada agrupades les dades el primer dels intervals de classe és [152.5,155.5) amb una freqüència igual a 2, un tractament discret de la variable continuarà considerant 154 com a valor mínim de la distribució ja que aquest és el valor central de l'anterior interval. Però un tractament continu considerara 153.25 com valor mínim d'aquesta distribució:

```
                153.25          154.75

        152.5                           155.5

                         154
```

El mateix podríem dir pel que fa referència al valor màxim de la distribució. Per tant, el rang de la distribució quedarà també afectat segons quin sigui el tractament que es doni a la variable numèrica.

b. Quartils.

El tractament discret d'una variable numèrica no agrupada en intervals de classe comporta que el valor d'un quartil - sigui quin sigui - o bé coincideixi amb una dada de la distribució (suposada ordenada) o bé sigui igual a la mitjana aritmètica de dues dades consecutives. Si la variable estigués agrupada en intervals, el procés que se seguiria seria el mateix, substituint les dades per les corresponents marques de classe.

Mentre que, si la variable és considerada de tipus continu, una vegada localitzat l'interval que conté el quartil que s'intenta calcular, la determinació d'aquest es realitzarà fent servir les dades de l'interval uniformement distribuïdes dins d'aquest.

Així, p.e., si una distribució consta de 50 dades i la freqüència acumulada dels intervals anteriors al [165.5,168.5) és 22 i la freqüència d'aquest és 10, un tractament discret de la variable que s'està manipulant considerarà la mediana (2on quartil) igual a la marca de classe 167 puix que les dades que ocupen els llocs 25è i 26è en la distribució s'identifiquen ambdues amb 167. Mentre que si la variable és tractada com a contínua la mediana serà igual a 166.40 semisuma dels dos valors centrals de la distribuió - 166.25 i 166.55 - i que ocupen els llocs 3er i 4rt dins l'anterior interval.

c. Desviació típica.

Aquest paràmetre també varia en funció del tipus de tractament que es doni a la variable. En general, la desviació típica d'una variable tractada com a contínua és més gran que la que resulta quan és considerada com a variable discreta. Les diferències entre ambdues desviacions són, en general, inapreciables.

Un altre aspecte que pot cridar l'atenció, és la importància que el programa atorga a la DIFERÈNCIA MÍNIMA entre valors no iguals d'una variable numèrica. En el fons, però, el que es té en compte és la PRECISIÓ amb què s'han realitzat les mesures de la variable numèrica. Així, per exemple, si les alçàries dels individus d'una població són precisades fins els cm, quan afirmem que una persona fa 172 cm d'estatura, estem dient pròpiament que la seva alçària real és un valor dins l'interval [171.5,172.5).

El programa té en compte aquesta diferència mínima o precisió a l'hora de:

- Agrupar les dades en intervals de classe.
  Així, no són permesos intervals de classe de longitud inferior a la precisió amb què s'han efectuat les mesures dels diferents valors de la variable. A més a més, si aquesta és tractada com a variable discreta, s'obliga a fer que la longitud dels intervals sigui múltiple de la precisió.

- Presentar el valor dels diferents paràmetres estadístics.

El programa, en funció de quina sigui la diferència mínima declarada, decideix quina ha d'ésser la precisió en la presentació dels paràmetres estadístics. Així, per exemple, si la diferència mínima és un nombre comprès entre 1 i 10, la precisió en la presentació de valors serà de l'ordre de 0.1, mentre que, si la diferència mínima és un valor que oscil.la entre 10 i 100, llavors la precisió en la presentació serà de l'ordre d'1 unitat. A aquest respecte, és important remarcar que, per tal d'evitar al màxim els errors d'arrodoniment que inevitablement realitza el programa a l'hora de realitzar els còmputs, és preferible operar amb valors enters que no pas amb valors decimals. Així, p.e., val més expressar les alçàries en centímetres que no pas en metres o decímetres. Per altra part,la limitació a 5 caràcters que el programa imposa als diferents valors que poden adquirir les variables, fa que sigui preferible operar amb valors que no utilitzin el punt decimal per estalviar, d'aquesta manera, un caràcter en l'escriptura d'aquells valors.

3.4 METODOLOGIA D'ÚS

Amb un conjunt de programes d'aquestes característiques, resulta difícil de donar indicacions sobre quina pot ésser la millor metodologia d'utilització ja que entenem que poden ésser utilitzats en el context de metodologies molt diverses. Per aquest motiu, ens limitarem a fer només uns quants suggeriments molt breus al respecte.

3.4.1 Utilització puntual dels programes

En un moment determinat i, normalment, fora del context de la classe de matemàtiques, aquests programes es poden utilitzar per a representar gràficament determinades distribucions estadístiques i calcular-ne els seus paràmetres més representatius.

Exemples:
- Estudi de la possible correlació existent entre les mitjanes de les temperatures mensuals màxima i mínima d'una localitat en el decurs d'un o més anys.
- Correlació entre pluviositat d'una comarca i nivell dels seus embassaments. Estudi comparatiu en el decurs d'un any.
- Alçària d'una planta en germinació en funció del temps.
- Diàmetre de les "pedres de riu" en funció del tram del riu on apareixen.

3.4.2 Utilització en el tractament estadístic d'una enquesta

Per algunes disciplines, com poden ser l'ètica, la geografia econòmica, la sociologia - dins el marc de la filosofia - ... , que tot sovint utilitzen la tècnica de les petites enquestes d'opinió a l'hora de presentar determinats temes, pot resultar interessant la utilització d'aquests programes encara que només sigui per obtenir un llistat general de les dades recollides i les diferents taules de freqüències. Per altra part, la utilització dels programes obligaria indirectament els alumnes a realitzar un disseny molt més acurat de les enquestes i a una inevitable codificació de les possibles respostes per tal de poder donar un tractament informàtic a les dades recollides.

3.4.3 Utilització com a suport en l'ensenyament de l'estadística

Resulta pràctica habitual que els alumnes s'iniciïn en el coneixement de l'estadística fent servir dades recollides per ells mateixos. La utilització d'un suport informàtic permet d'ampliar el volum de dades a tractar i també el nombre de variables.

Així, per exemple, mentre que a classe es treballen "manualment" les dades del grup, simultàniament hom pot anar comparant els resultats que es vagin obtenint amb els relatius a d'altres grups de manera que no hagin d'ésser calculats "manualment" sinó que siguin proporcionats pel programa.

L'arxiu TERCERBC.DDG que acompanya els programes recull tota una sèrie de dades relatives al sexe, l'alçada, la distància entre les puntes de les mans amb els braços, ... relatives als 30 alumnes d'un grup de 3er curs.

3.4.4 Utilització com a reforç de la comprensió

Ja s'ha parlat amb anterioritat d'aquesta possibilitat basada fonamentalment en les facilitats gràfiques que proporciona el programa.

4 APÈNDIX

A continuació s'inclouen els llistats dels continguts dels diferents arxius que acompanyen els programes. Les dades que hi figuren són dades reals o han estat extretes de diferent bibliografia.

4.1 BREU DESCRIPCIÓ DE CADA ARXIU

4.1.1 TERCERBC.DDG

- Dades recollides en un grup de 3er curs de B.U.P. de l'I.B. "SALVADOR ESPRIU", de SALT.

  SEXE = sexe
  ALÇA = alçària (precisió 1 cm)
  PMANS = distància entre les puntes de les mans amb els braços oberts (precisió 1 cm)
  PES = pes (precisió 1 kg)
  NSAB = número de sabata
  NGERM = nombre de germans (inclòs l'alumne entrevistat)
  DINST = distància a l'institut del domicili familiar (precisió 100 m). Als alumnes residents a més de 3000 m se'ls ha adjudicat a tots una distància de 6000 m.
  TEMPS = temps per anar del domicili familiar a l'institut (precisió 1 minut).
  TRANS = mitjà de transport
  ESPOR = esport preferit (espectacle o pràctica)
  COLOR = color preferit
  MATEM = qualificació matemàtiques 2on curs (juny o setembre)

4.1.2 OUS1.DD2

Pes i diàmetre de 5 dotzenes d'ous de qualitat "extra" d'una marca "A" . A l'etiqueta hi figurava que el pes dels ous estava comprés entre 60 i 67 g . El pes s'ha presentat fins als dg i el diàmetre fins als mm.

4.1.3 OUS2.DD1

Pes de 5 dotzenes d'ous de qualitat "extra" d'una marca "B". A l'etiqueta hi figurava "pes entre 60 i 65 g ".

4.1.4 MEITAT.DD4

Quatre persones A , B , C i D , a "cop d'ull", divideixen per la meitat 12 segments dibuixats sobre sengles fulls de paper. Repeteixen l'experiment 12 vegades consecutives fent servir sempre segments de la mateixa longitud. Les dades de l'arxiu recullen les longituds de les "meitats esquerres" de les 12 tires dividides per a cada un dels quatre individus.

Font: Kanji, G. (1982) "Developing materials for 16-19 year old students"

4.1.5 TEMPWAME.DD2

Temperatures mitjanes mensuals - en el decurs d'un mateix any - a les ciutats de Washington (TWASH) i Melbourne (TMELB). Precisió: 1º F .

Font: A.N. Matsumoto, "Correlation, Junior Varsity Style", Teaching Statistics and Probability, NCTM, 1981.

4.1.6 ALPARFIL.DD2

Alçària de pare i fill. Precisió: 1 cm

Font: A.N. Matsumoto, "Correlation, Junior Varsity Style", Teaching Statistics and Probability, NCTM, 1981.

4.1.7 ENCRPCAS.DF1

Taula de freqüències de la llargada - en lletres - de totes les paraules utilitzades en 6 entreteniments de "mots encreuats" (en castellà) del diari "El País" durant el mes de novembre de 1986.

4.1.8 COTXESAN.DF1

Durant mitja hora, i a intervals de 30 s, es varen comptabilitzar el nombre de cotxes que entraven a la població de Salt procedents d'Anglès a la 1 del migdia. L'arxiu conté el nombre d'intervals temporals que varen enregistrar una freqüència d'entrada de 0 cotxes, 1 cotxe, ... , fins a 8 cotxes.

4.1.9 COTXAUBA.DF1

El mateix tipus d'arxiu que l'anterior comptabilitzant els cotxes que circulaven per l'autopista procedents de Barcelona, a l'altura de la sortida de Girona-sud.

4.1.10 COTXAUFR.DF1

Igual que l'arxiu anterior comptabilitzant, però, els cotxes procedents de França.

4.1.11 <u>BOMBES.DF1</u>

Imaginant per un moment la ciutat de Londres dividida en àrees de 0.25 km^2 d'extensió, l'arxiu recull el nombre d'aquestes àrees que varen rebre 0 , 1 , 2 , ... fins a 8 impactes de bomba durant la IIa Guerra Mundial.

Font: Clarke, R.D. "An Application of the Poisson Distribution", Journal of the Institute of Actuaries, 1946.

4.1.12 <u>PARTICUL.DF1</u>

Nombre de partícules α emeses per una substància radioactiva durant 2608 períodes de 7.5 segons cada un. L'arxiu recull els resultats obtinguts per Rutherford i Geiger l'any 1930.

Font: Fisz, M. "Probability Theory and Mathematical Statistics", New York, John Wiley & Sons, 1963.

4.1.13 <u>ACCIEDAT.DF2</u>

Nombre d'accidents d'automòbil en funció de l'edat.

Font: Murray R. Spiegel, "Estadística", Libros McGraw-Hill, 1970.

4.1.14 <u>FISMAT.DF2</u>

Qualificacions en Física i Matemàtiques d'un grup de 100 alumnes. Cada qualificació es barema sobre 100 punts.

Font: Murray R. Spiegel, "Estadística", Libros McGraw-Hill, 1970.

4.2 <u>EXEMPLES D'EXECUCIÓ</u>

Aquest apèndix finalitza presentant dos exemples d'execució dels programes.

El primer exemple està dedicat a l'estudi de les variables ALÇA i MATEM de l'arxiu TERCERBC.DDG, i el segon a l'anàlisi de les variables FISIC i MATEM de l'arxiu FISMAT.DF2 .

Nom arxiu ...: TERCERBC.DDG

Breu descripció ...: DADES ALUMNES 3r-B I.B."S.ESPRIU" 86-87

Relació de variables:

	NOM	TIPUS	UNITAT
1	SEXE	ALF	
2	ALÇA	NUM	cm.
3	PMANS	NUM	cm.
4	PES	NUM	kg.
5	NSAB	NUM	
6	NGERM	NUM	germ
7	DINST	NUM	m.
8	TEMPS	NUM	min.
9	TRANS	ALF	
10	ESPOR	ALF	
11	COLOR	ALF	
12	MATEM	ALF	

REG	SEXE	ALÇA	PMANS	PES	NSAB	NGERM	DINST	TEMPS	TRANS	ESPOR	COLOR	MATEM
1	D	163	161	55	36	3	500	7	A_PEU	BASQ	BLAU	SUF
2	D	166	166	56	39	6	1100	18	A_PEU	HAND	BLAU	BE
3	D	158	157	43	36	3	500	10	A_PEU	NATA	BLAU	SUF
4	D	162	160	51	37	4	1100	16	A_PEU	CAP	BLAU	NOT
5	D	161	160	60	38	1	800	12	A_PEU	BASQ	BLANC	BE
6	D	161	157	61	39	9	500	10	A_PEU	BASQ	VERM	SUF
7	D	162	166	44	36	3	6000	10	COTXE	PATI	BLAU	BE
8	D	166	161	64	38	1	1400	24	A_PEU	ALTR	HAND	SUF
9	D	159	155	52	37	1	400	7	A_PEU	FUTB	BASQ	BE
10	D	155	150	50	35	4	1400	22	A_PEU	BASQ	VERM	BE
11	D	167	167	60	37	1	6000	10	BUS	BASQ	VERM	SUF
12	D	168	169	64	40	2	500	12	A_PEU	CAP	BLAU	BE
13	H	162	164	50	40	6	200	2	A_PEU	FUTB	VERM	BE
14	H	156	163	47	39	2	1100	20	A_PEU	ESQUI	BLAU	SUF
15	D	161	162	51	37	2	1100	16	A_PEU	ALTR	VERD	NOT
16	D	160	163	63	38	3	1400	25	A_PEU	HAND	VERM	SUF
17	H	177	175	58	43	1	1100	23	A_PEU	BASQ	BLAU	BE
18	H	177	181	71	41	2	2000	6	MOTO	ALTR	BLAU	SUF
19	H	166	169	59	40	2	500	10	A_PEU	BASQ	VERD	SUF
20	H	171	181	77	44	3	800	16	A_PEU	BASQ	BLAU	SUF
21	D	164	165	64	36	2	800	12	A_PEU	BASQ	BLAU	SUF
22	D	159	150	53	37	4	800	15	A_PEU	BASQ	NEGRE	SUF
23	D	165	162	50	38	2	800	8	A_PEU	BASQ	VERM	SUF
24	D	153	155	45	36	4	800	8	A_PEU	PATI	VERD	NOT
25	D	163	161	53	38	2	6000	10	BUS	HAND	BLAU	BE
26	D	169	163	65	38	2	200	5	A_PEU	NATA	BLAU	BE
27	D	157	162	47	38	1	800	12	A_PEU	NATA	NEGRE	SUF
28	D	171	172	70	40	3	800	10	A_PEU	NATA	VERM	SUF
29	D	168	166	60	38	2	6000	10	BUS	CAP	NEGRE	NOT
30	D	161	159	47	37	3	6000	10	BUS	CAP	BLAU	NOT

Nom arxiu ...: OUS1.DD2

Breu descripció ...: PES i DIAMETRE OUS (EXTRA) MARCA "A"

Relació de variables:

	NOM	TIPUS	UNITAT
1	PES	NUM	dg
2	DIAM	NUM	mm

REG	PES	DIAM	REG	PES	DIAM	REG	PES	DIAM
1	625	57	21	632	59	41	632	59
2	616	57	22	642	58	42	609	58
3	627	58	23	617	60	43	627	57
4	647	58	24	637	57	44	657	60
5	642	60	25	638	60	45	634	60
6	634	59	26	654	59	46	652	59
7	629	57	27	627	56	47	657	57
8	634	61	28	622	58	48	657	58
9	619	61	29	650	61	49	623	57
10	613	59	30	622	57	50	622	61
11	635	58	31	622	58	51	662	62
12	660	60	32	642	58	52	657	60
13	624	57	33	642	60	53	637	56
14	607	56	34	619	58	54	609	60
15	665	59	35	627	57	55	632	57
16	612	59	36	640	58	56	627	59
17	613	57	37	607	55	57	662	56
18	637	60	38	630	58	58	647	61
19	659	57	39	610	58	59	642	59
20	657	59	40	600	58	60	619	59

Nom arxiu ...: OUS2.DD1

Breu descripció ...: PES OUS (EXTRA) MARCA "B"

Relació de variables:

	NOM	TIPUS	UNITAT
1	PES	NUM	dg

REG	PES	REG	PES	REG	PES
1	579	21	617	41	628
2	609	22	636	42	622
3	609	23	584	43	621
4	626	24	582	44	604
5	596	25	604	45	598
6	597	26	624	46	574
7	632	27	590	47	589
8	596	28	597	48	623
9	602	29	610	49	617
10	623	30	625	50	487
11	594	31	624	51	617
12	603	32	624	52	594
13	621	33	615	53	607
14	634	34	605	54	616
15	622	35	603	55	594
16	614	36	603	56	604
17	594	37	596	57	603
18	629	38	584	58	599
19	611	39	599	59	639
20	604	40	[illegible]9	60	600

Nom arxiu ...: MEITAT.DD4

Breu descripció ...: MEITAT A ULL DE 12 LONGITUDS (4 indiv.)

Relació de variables:

	NOM	TIPUS	UNITAT
1	MESA	NUM	mm
2	MESB	NUM	mm
3	MESC	NUM	mm
4	MESD	NUM	mm

REG	MESA	MESB	MESC	MESD
1	304	332	326	311
2	316	323	314	310
3	310	328	320	315
4	301	335	318	316
5	312	326	323	312
6	307	320	327	298
7	313	324	319	304
8	307	332	325	304
9	316	319	322	326
10	318	326	319	320
11	319	341	318	314
12	327	303	322	313

Nom arxiu ...: FABRICA.DD1

Breu descripció ...: PESOS MOSTRA PECES D'UNA FABRICA

Relació de variables:

	NOM	TIPUS	UNITAT
1	PES	NUM	g

REG	PES						
1	443	26	469	51	426	76	449
2	455	27	440	52	466	77	430
3	436	28	450	53	452	78	461
4	442	29	441	54	456	79	467
5	441	30	442	55	440	80	470
6	440	31	451	56	428	81	459
7	462	32	460	57	431	82	461
8	468	33	452	58	459	83	450
9	449	34	443	59	456	84	430
10	450	35	452	60	477	85	468
11	470	36	453	61	442	86	460
12	446	37	454	62	462	87	454
13	472	38	427	63	470	88	466
14	470	39	460	64	445	89	442
15	465	40	446	65	450	90	441
16	469	41	471	66	448	91	424
17	441	42	466	67	455	92	465
18	449	43	449	68	438	93	459
19	429	44	456	69	447	94	437
20	462	45	429	70	449	95	464
21	470	46	432	71	459	96	464
22	451	47	460	72	458	97	465
23	431	48	459	73	458	98	430
24	432	49	458	74	454	99	475
25	449	50	476	75	437	100	462

Nom arxiu ...: TEMPWAME.DD2

Breu descripció ...: TEMPERATURES MITJ.MENS. WASHING.-MELBOUR

Relació de variables:

	NOM	TIPUS	UNITAT
1	TWASH	NUM	°F
2	TMELB	NUM	°F

REG	TWASH	TMELB
1	44	78
2	45	78
3	55	75
4	65	68
5	75	62
6	83	57
7	86	56
8	84	59
9	78	63
10	67	67
11	56	71
12	45	75

Nom arxiu ...: ALPARFIL.DD2

Breu descripció ...: ALÇADA PARE - FILL

Relació de variables:

	NOM	TIPUS	UNITAT
1	APARE	NUM	cm
2	AFILL	NUM	cm

REG	APARE	AFILL
1	178	175
2	177	175
3	169	173
4	176	170
5	163	163
6	180	179
7	173	163
8	173	169
9	188	193
10	170	188
11	175	170
12	168	165
13	201	193
14	165	163
15	173	180
16	173	173
17	160	152

Nom arxiu ...: ENCRPCAS.DF1

Breu descripció ...: LLARGADA MOTS ENCREUATS (EL PAIS-cast.)

Relació de variables:

	NOM	TIPUS	UNITAT	AGRUP	DIFER	V.MIN
1	LONG	NUM	11	NO	1	1

REG	VALOR 1a	VALOR 2a	FREQUENCIA
1	1		11
2	2		18
3	3		13
4	4		22
5	5		16
6	6		13
7	7		16
8	8		3
9	9		6
10	10		1
11	11		1

Nom arxiu ...: COTXESAN.DF1

Breu descripció ...: COTXES ENTREN A SALT (d'Anglès) CADA 30s

Relació de variables:

	NOM	TIPUS	UNITAT	AGRUP	DIFER	V.MIN
1	NCOTX	NUM	cotx	NO	1	0

REG	VALOR 1a	VALOR 2a	FREQUENCIA
1	0		4
2	1		11
3	2		16
4	3		9
5	4		11
6	5		2
7	6		2
8	7		2
9	8		3

Nom arxiu ...: COTXAUBA.DF1

Breu descripció ...: COTXES AUTOPISTA (de Barna) cada 30 s.

Relació de variables:

	NOM	TIPUS	UNITAT	AGRUP	DIFER	V.MIN
1	NCOTX	NUM	cotx	NO	1	0

REG	VALOR 1a	VALOR 2a	FREQUENCIA
1	0		4
2	1		8
3	2		17
4	3		20
5	4		8
6	5		6
7	6		2
8	7		2
9	8		1
10	9		1

Nom arxiu ...: COTXAUFR.DF1

Breu descripció ...: COTXES AUTOPISTA (de França) cada 30 s.

Relació de variables:

	NOM	TIPUS	UNITAT	AGRUP	DIFER	V.MIN
1	NCOTX	NUM	cotx	NO	1	0

REG	VALOR 1a	VALOR 2a	FREQUENCIA
1	0		2
2	1		9
3	2		12
4	3		11
5	4		11
6	5		8
7	6		3
8	7		3
9	8		1

Nom arxiu ...: BOMBES.DF1

Breu descripció ...: NOMBRE BOMBES SOBRE AREES 1/4 km² LONDON

Relació de variables:

	NOM	TIPUS	UNITAT	AGRUP	DIFER	V.MIN
1	NBOMB	NUM	bomb	NO	1	0

REG	VALOR 1a	VALOR 2a	FREQUENCIA
1	0		229
2	1		211
3	2		93
4	3		35
5	4		7
6	5		0
7	6		0
8	7		1
9	8		0

Nom arxiu ...: PARTICUL.DF1

Breu descripció ...: NOMBRE PARTICLES α EMESES CADA 7.5 s

Relació de variables:

	NOM	TIPUS	UNITAT	AGRUP	DIFER	V.MIN
1	NPART	NUM	α	NO	1	0

REG	VALOR 1a	VALOR 2a	FREQUENCIA
1	0		57
2	1		203
3	2		383
4	3		525
5	4		532
6	5		408
7	6		273
8	7		139
9	8		45
10	9		27
11	10		16

Nom a-xiu ...: ACCIEDAT.DF2

Breu descripció ...: NOMBRE ACCIDENTS EN FUNCIO EDAT

Relació de variables:

	NOM	TIPUS	UNITAT	AGRUP	DIFER	V.MIN
1	EDAT	NUM	anys	SI	10	25
2	ACCID	NUM	acc.	SI	1	0

REG	VALOR 1a	VALOR 2a	FREQUENCIA
1	25	0	748
2	25	1	74
3	25	2	31
4	25	3	9
5	35	0	821
6	35	1	60
7	35	2	25
8	35	3	10
9	45	0	786
10	45	1	51
11	45	2	22
12	45	3	6
13	55	0	720
14	55	1	66
15	55	2	16
16	55	3	5
17	65	0	672
18	65	1	50
19	65	2	15
20	65	3	7

CORRECCIO DE DADES ARXIUS DISTRIBUCIO DE FREQUENCIES

ACCIEDAT.DF2

	EDAT*				
ACCID	25	35	45	55	65
0	748	821	786	720	672
1	74	60	51	66	50
2	31	25	22	16	15
3	9	10	6	5	7

Nom arxiu ...: FISMAT.DF2

Breu descripció ...: NOTES FISICA-MATEMATIQUES (SOBRE 100 p)

Relació de variables:

	NOM	TIPUS	UNITAT	AGRUP	DIFER	V.MIN
1	MATEM	NUM	p	SI	10	44.5
2	FISIC	NUM	p	SI	10	44.5

REG	VALOR 1a	VALOR 2a	FREQUENCIA
1	44.5	44.5	3
2	44.5	54.5	3
3	44.5	64.5	1
4	44.5	74.5	0
5	44.5	84.5	0
6	44.5	94.5	0
7	54.5	44.5	5
8	54.5	54.5	6
9	54.5	64.5	4
10	54.5	74.5	0
11	54.5	84.5	0
12	54.5	94.5	0
13	64.5	44.5	4
14	64.5	54.5	6
15	64.5	64.5	9
16	64.5	74.5	5
17	64.5	84.5	1
18	64.5	94.5	0
19	74.5	44.5	0
20	74.5	54.5	2
21	74.5	64.5	5
22	74.5	74.5	10
23	74.5	84.5	4
24	74.5	94.5	2
25	84.5	44.5	0
26	84.5	54.5	0
27	84.5	64.5	2
28	84.5	74.5	8
29	84.5	84.5	6
30	84.5	94.5	4
31	94.5	44.5	0
32	94.5	54.5	0
33	94.5	64.5	0
34	94.5	74.5	1
35	94.5	84.5	5
36	94.5	94.5	4

CORRECCIO DE DADES ARXIUS DISTRIBUCIO DE FREQUENCIES FISMAT.DF2

FISIC	MATEM 44.5_	54.5_	64.5_	74.5_	84.5_	94.5_
44.5	3	5	4	0	0	0
54.5	3	6	6	2	0	0
64.5	1	4	9	5	2	0
74.5	0	0	5	10	8	1
84.5	0	0	1	4	6	5
94.5	0	0	0	2	4	4

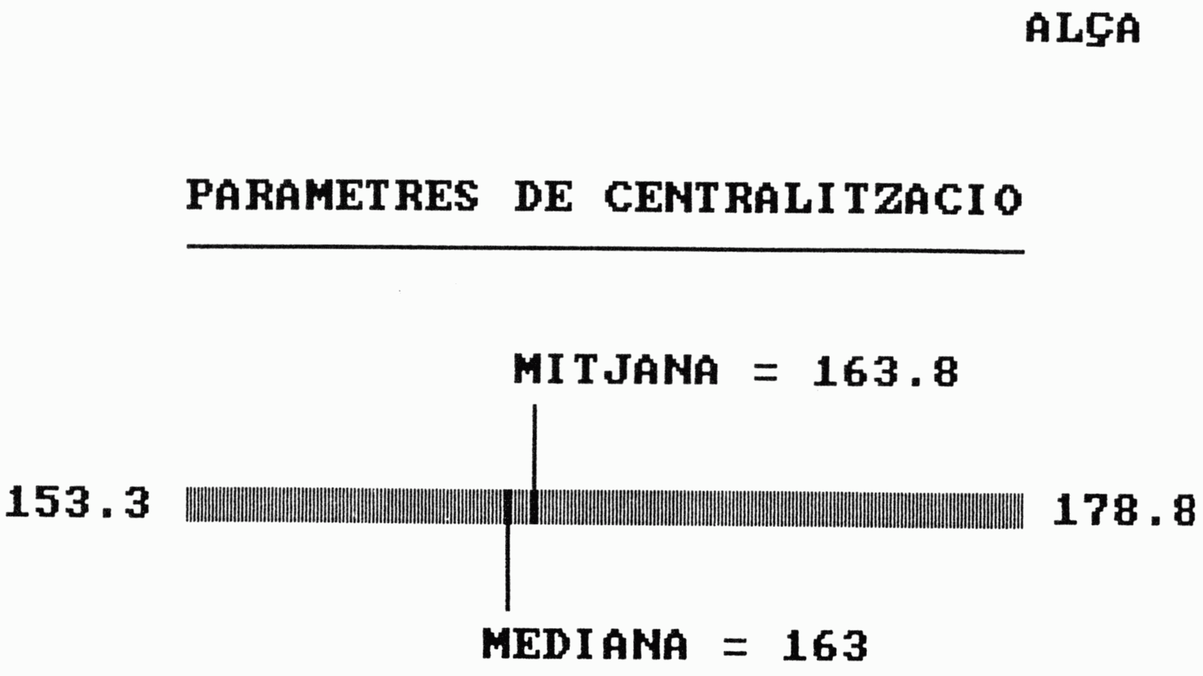
<┘ = CORREGIR   F1=ACABAR OPCIO   F9=INTERRUPCIO PROGR.

TAULA DE FREQUENCIES DE LA VARIABLE ALÇA

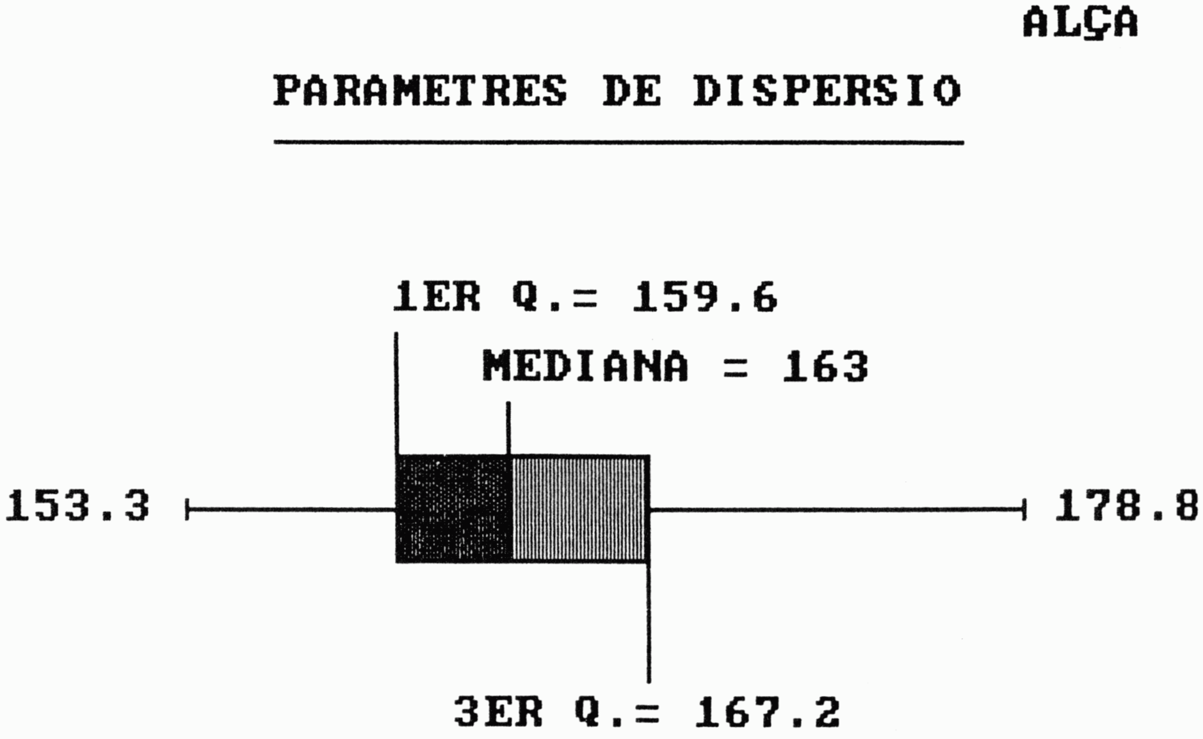
	Intervals	x	F	%F	FA	%FA
1	[152.5 - 155.5)	154	2	6.7	2	6.7
2	[155.5 - 158.5)	157	3	10.0	5	16.7
3	[158.5 - 161.5)	160	7	23.3	12	40.0
4	[161.5 - 164.5)	163	6	20.0	18	60.0
5	[164.5 - 167.5)	166	5	16.7	23	76.7
6	[167.5 - 170.5)	169	3	10.0	26	86.7
7	[170.5 - 173.5)	172	2	6.7	28	93.3
8	[173.5 - 176.5)	175	0	0.0	28	93.3
9	[176.5 - 179.5)	178	2	6.7	30	100.0

30

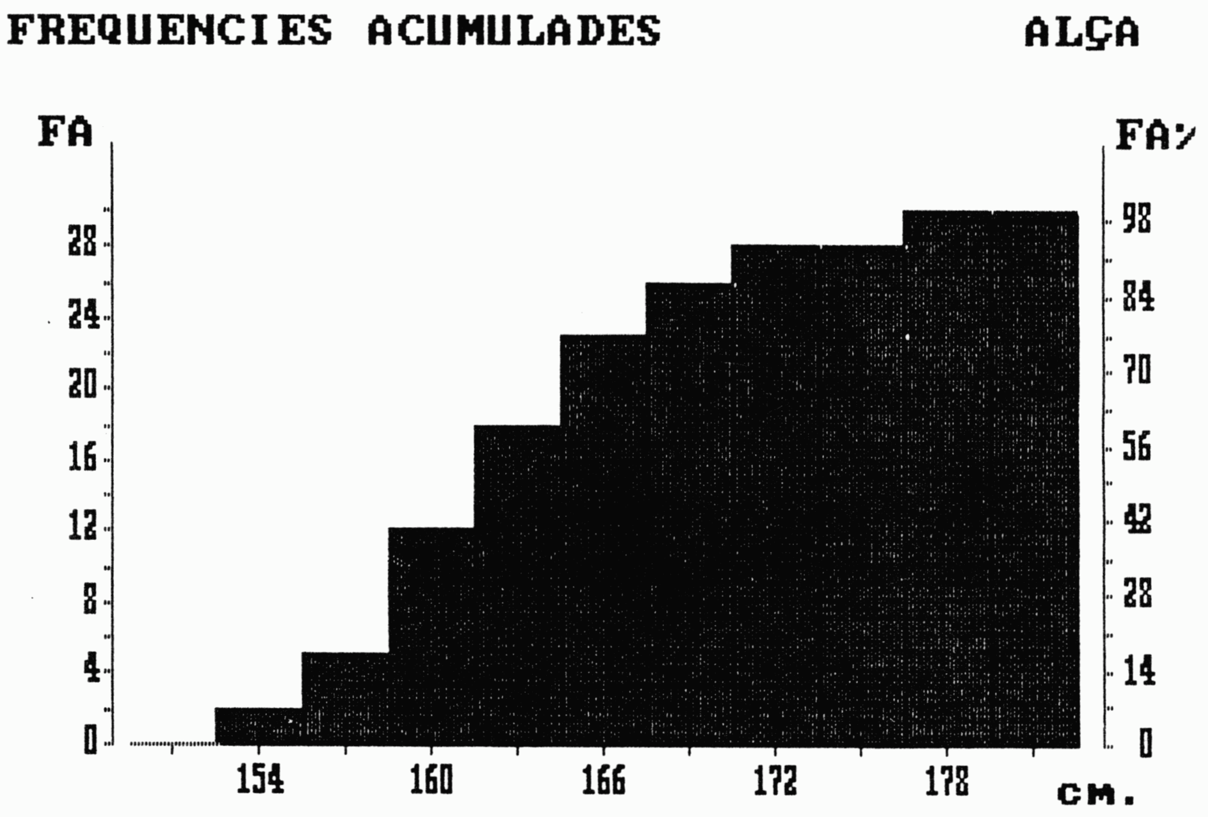
POLSEU L'ESPAIADOR   F9=INTERROMPRE PROGR.



ALÇA
PARAMETRES DE CENTRALITZACIO
MITJANA = 163.8
153.3
178.8
MEDIANA = 163



ALÇA
PARAMETRES DE DISPERSIO
1ER Q.= 159.6
MEDIANA = 163
153.3
178.8
3ER Q.= 167.2



TAULA DE FREQUENCIES DE LA VARIABLE MATEM

	Valor	F	%F
1	SUF	15	50.0
2	BE	10	33.3
3	NOT	5	16.7
		30	

DIAGRAMA DE SECTORS **MATEM**

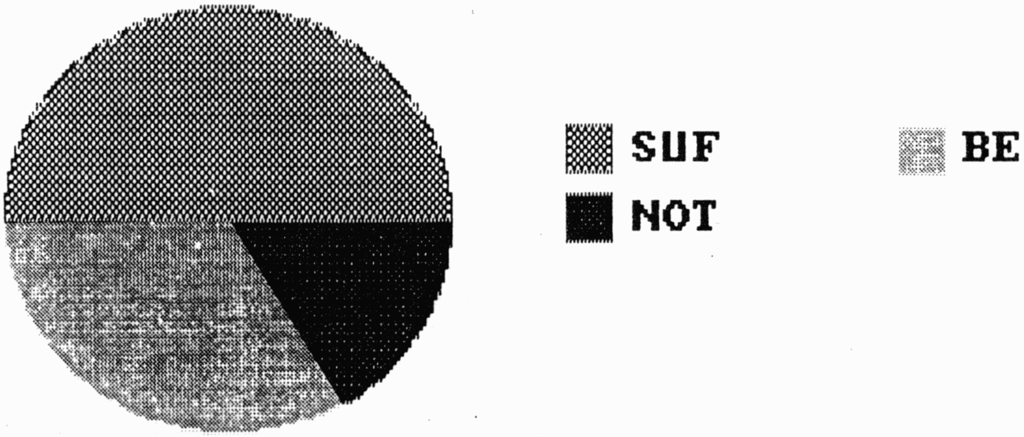
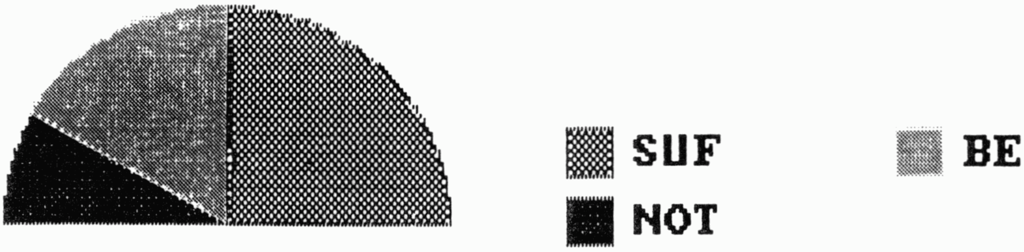
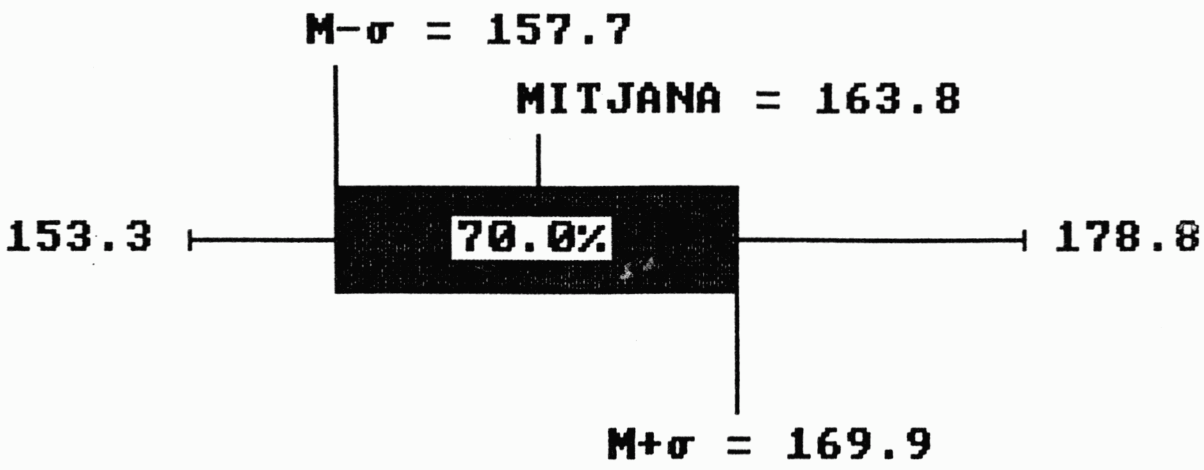


DIAGRAMA DE SECTORS **MATEM**

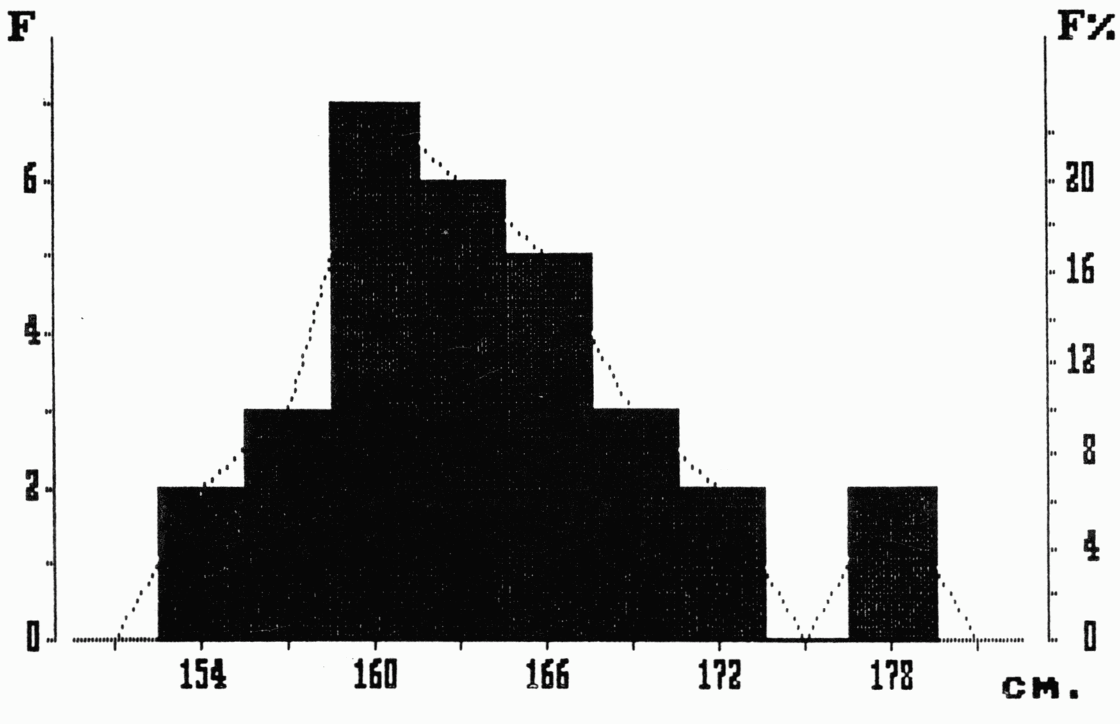


ALÇA

PARAMETRES DE DISPERSIO



FREQUENCIES ABSOLUTES **ALÇA**



TAULA DE FREQUENCIES ABSOLUTES

FISIC	MATEM 44.5	54.5	64.5	74.5	84.5	94.5	
_44.5	3	5	4	0	0	0	12
_54.5	3	6	6	2	0	0	17
_64.5	1	4	9	5	2	0	21
_74.5	0	0	5	10	8	1	24
_84.5	0	0	1	4	6	5	16
_94.5	0	0	0	2	4	4	10
	7	15	25	23	20	10	100

POLSEU L'ESPAIADOR F9=INTERROMPRE PROGR.

TAULA DE FREQUENCIES PERCENTUALS

FISIC	MATEM 44.5	54.5	64.5	74.5	84.5	94.5	
_44.5	3.0	5.0	4.0	0.0	0.0	0.0	12.0
_54.5	3.0	6.0	6.0	2.0	0.0	0.0	17.0
_64.5	1.0	4.0	9.0	5.0	2.0	0.0	21.0
_74.5	0.0	0.0	5.0	10.0	8.0	1.0	24.0
_84.5	0.0	0.0	1.0	4.0	6.0	5.0	16.0
_94.5	0.0	0.0	0.0	2.0	4.0	4.0	10.0
	7.0	15.0	25.0	23.0	20.0	10.0	100.0

POLSEU L'ESPAIADOR F9=INTERROMPRE PROGR.

TAULA DE FREQUENCIES PERCENTUALAS PER COLUMNES

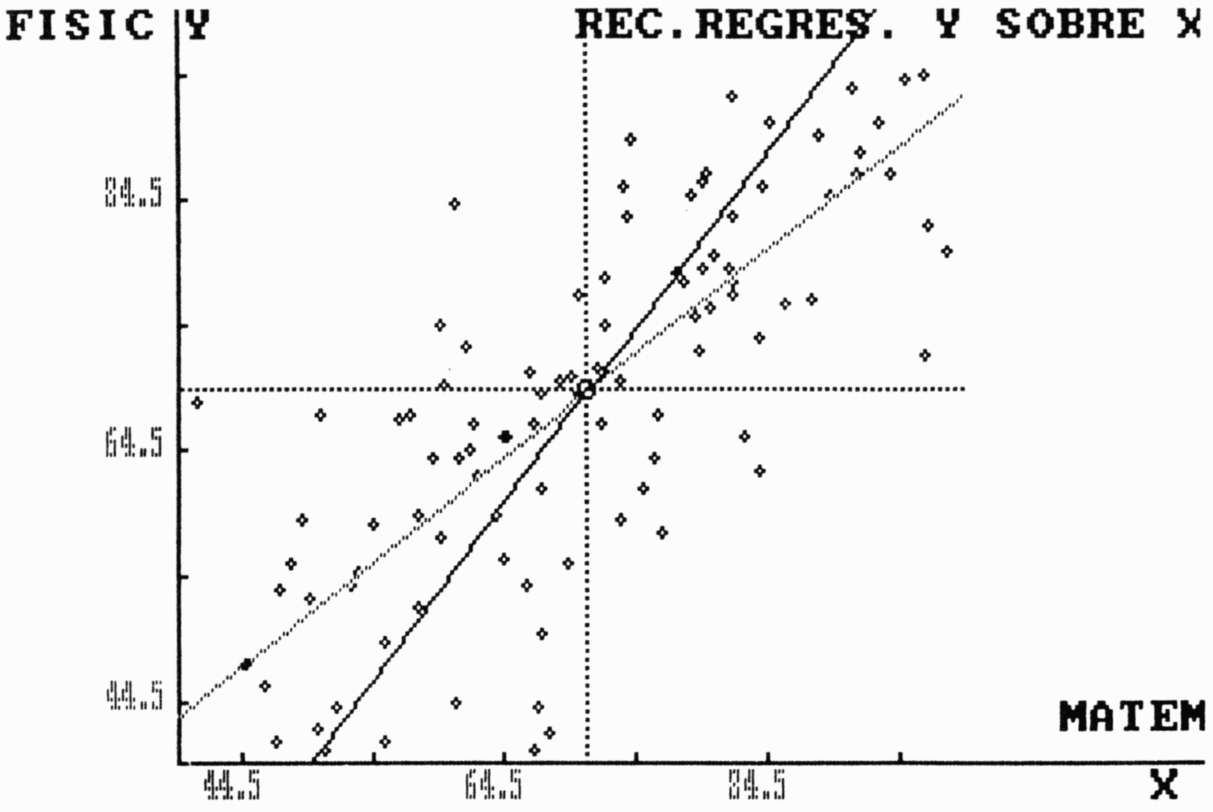
FISIC	MATEM 44.5	54.5	64.5	74.5	84.5	94.5	
_44.5	42.9	33.3	16.0	0.0	0.0	0.0	12.0
_54.5	42.9	40.0	24.0	8.7	0.0	0.0	17.0
_64.5	14.3	26.7	36.0	21.7	10.0	0.0	21.0
_74.5	0.0	0.0	20.0	43.5	40.0	10.0	24.0
_84.5	0.0	0.0	4.0	17.4	30.0	50.0	16.0
_94.5	0.0	0.0	0.0	8.7	20.0	40.0	10.0
	100.0	100.0	100.0	100.0	100.0	100.0	100.0

POLSEU L'ESPAIADOR F9=INTERROMPRE PROGR.

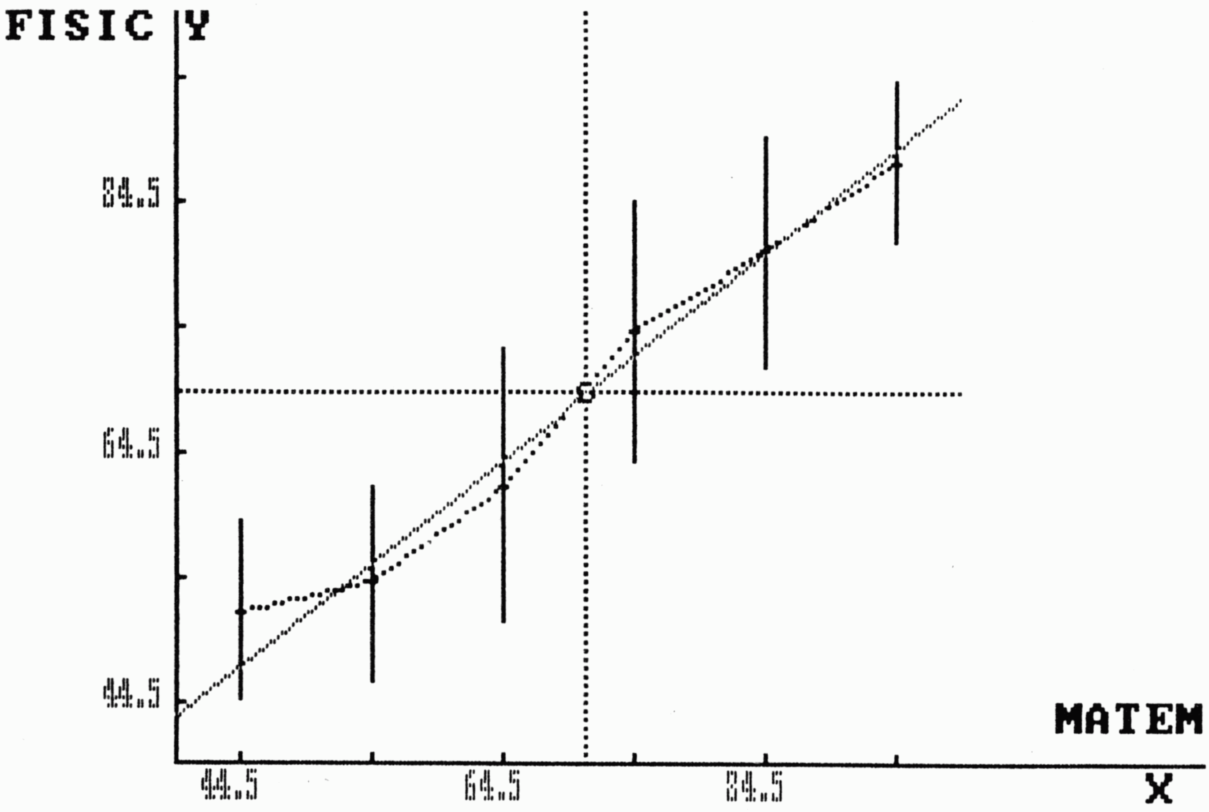
TAULA DE FREQUENCIES PERCENTUALS PER FILES

FISIC	MATEM 44.5	54.5	64.5	74.5	84.5	94.5	
_44.5	25.0	41.7	33.3	0.0	0.0	0.0	100.0
_54.5	17.6	35.3	35.3	11.8	0.0	0.0	100.0
_64.5	4.8	19.0	42.9	23.8	9.5	0.0	100.0
_74.5	0.0	0.0	20.8	41.7	33.3	4.2	100.0
_84.5	0.0	0.0	6.3	25.0	37.5	31.3	100.0
_94.5	0.0	0.0	0.0	20.0	40.0	40.0	100.0
	7.0	15.0	25.0	23.0	20.0	10.0	100.0

POLSEU L'ESPAIADOR F9=INTERROMPRE PROGR.

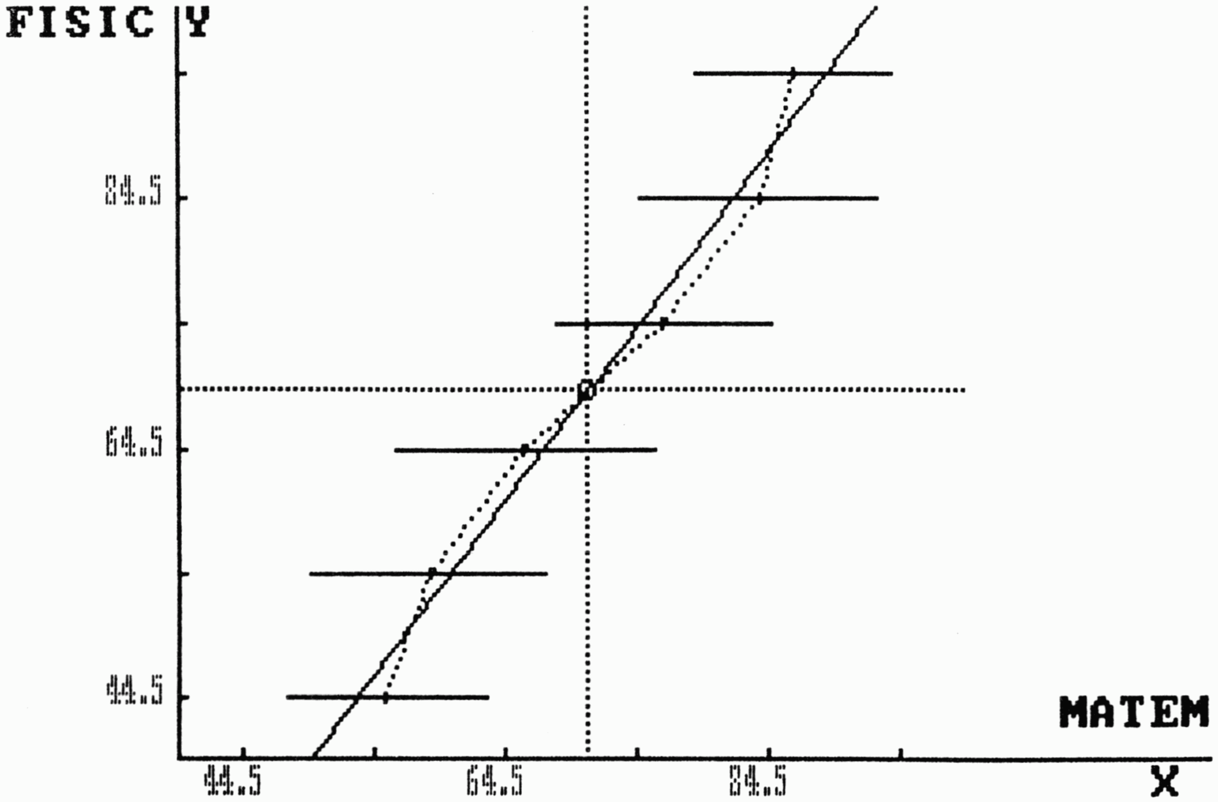


FISIC
Y
REC.REGRES. Y SOBRE X
84.5
64.5
44.5
MATEM
44.5
64.5
84.5
X



M i σ DE CADA COLUMNA

FISIC	MATEM 44.5	54.5	64.5	74.5	84.5	94.5	TOTAL
M	52	54	62	74	81	88	69
σ	7	8	11	10	9	6	15



M i σ DE CADA FILA

FISIC	MATEM M	MATEM σ
_44.5	55	8
_54.5	59	9
_64.5	66	10
_74.5	77	8
_84.5	84	9
_94.5	87	7
TOTAL	71	14